INVERSE METHODS AND MODELLING

A thesis

presented for the Degree of

Doctor of Philosophy in Electrical Engineering

in the

University of Canterbury,

Christchurch, New Zealand

by

R.P. MILLANE

University of Canterbury

1981

ABSTRACT

Three applications of inverse methods are considered.

The theoretical bases of the major inverse scattering techniques used to image the interior of penetrable bodies are reviewed.

X-ray crystallography is used to image the molecular structure of crystals. Two models of the DNA molecule are analysed on the basis of the X-ray diffraction data. The two models are the widely accepted "double helix" and the recently proposed "side-by-side". It is shown that the side-by-side model fits the diffraction data at least as well as the double helix. A stereochemical analysis shows that the side-by-side model is also stereochemically acceptable.

Current methods for imaging regions of variable refractive index are useful only for weak inhomogeneities. A time domain method for imaging one-dimensional regions of arbitrary variation in refractive index is outlined. It is shown that this method may be applied to branched transmission line networks. These techniques are applied to both simulated and measured data.

Physiological and clinical aspects of cardiac arrhythmias are reviewed.

A modelling approach to the inverse problem of cardiac arrhythmia diagnosis is outlined. The important variables describing cardiac conduction are identified and used as the basis of an interactive computer model to assist in arrhythmia diagnosis. It is shown that the model can simulate realistic quantitative rhythms. Methods of processing patients' clinical data to identify the model parameters are described. Examples of the use of the model to simulate two patients' arrhythmias are presented.

ACKNOWLEDGEMENTS

I sincerely thank my supervisor, Professor Richard Bates, for his continual encouragement and guidance throughout the course of the work reported in this thesis.

Special thanks are also due to Dr Gordon Rodley and Phil Bones for their considerable assistance with the work reported in chapters 2 and 5 respectively.

My colleagues from the Electrical Engineering Department, in particular Dr Graeme McKinnon, Dr Patrick Heffernan, Brent Robinson, Andrew Seagar and Richard Fright, all provided companionship and assistance.

I would also like to thank Dr Allan McKinnon, Robin McNeil, Adrian Coronno, Gillian Rodley, Marie Flewellen, Dr Andy Maslowski and Dr Hamid Ikram, all of whom have made a contribution to this thesis.

I am grateful to my family, Stephanie Morris, Kathy McSweeney and my friends and flatmates over the years for their support and for helping to keep me sane.

Finally, I am grateful for the assistance of the New Zealand University Grants Committee, the New Zealand Medical Research Council, the National Heart Foundation of New Zealand, and the New Zealand Post Office.

PREFACE

Direct observation of a physical system is often either undesirable, impracticable or impossible. In such a circumstance one must resort to indirect measurement schemes which involve the measurement of quantities related to those required, rather than those quantities themselves. These indirect measurement methods often fall within the realm of inverse problems. Inverse problems usually involve the determination of the cause of a phenomenon, given a measurement of its effect. Direct problems, on the other hand, generally involve the determination of an effect, given the cause. Inverse problems are, in general, significantly less well understood than the corresponding direct problems.

This thesis is concerned with three applications of inverse methods. Two of these (chapters 1 to 3) fall within the realm of "inverse scattering" which involves the determination of some of the properties of an object from measurement of radiation scattered by it. The third application (chapters 4 and 5) is concerned with physiological modelling to assist in the inverse problem of cardiac arrhythmia diagnosis. Review material is presented in chapters 1 and 4, and original work is reported in chapters 2, 3 and 5.

Chapter 1 contains most of the theoretical preliminaries required for chapters 2 and 3, and reviews the major inverse scattering techniques for imaging the interior of penetrable bodies. The chapter begins with derivations of the equations describing electromagnetic and acoustic waves in unbounded space and on guiding structures. This is followed by a description of the major inverse scattering techniques used in imaging, crystallography, echo-location, sounding and profile inversion.

X-ray crystallography is an inverse scattering technique used to image the molecular structure of crystals from measurement of their X-ray diffraction patterns. However structure determinations are often not unique because the phase of the diffraction pattern cannot be measured and one often measures incomplete data. These ambiguities can sometimes be resolved by making use of independent stereochemical (geometrical factors related to atomic positions) information. For many biological molecules, however, the ambiguity remains and the best that can be done is to build a trial model for the structure, calculate the diffraction pattern and

see how well it agrees with that observed. Adjustments are made to the model in an attempt to reduce the discrepancy between the calculated and observed diffraction patterns below an acceptable level. Chapter 2 describes the comparison of a new model for DNA (proposed by Dr Gordon Rodley of the Chemistry department of this University) with the accepted double helix model in terms of the observed X-ray diffraction data. A stereochemical assessment of Rodley's model is also described.

Most practical inverse scattering techniques used to image regions of variable refractive index provide useful results only if there is no need to go beyond simple echo-location principles. These methods are unsuitable for imaging inhomogeneities which are either strong or are of spatial extent which is large in comparison to the wavelength of the radiation. In chapter 3 a method, formulated in the time domain, for imaging plane stratified regions with arbitrary variations in refractive index is described. This method is inherently unstable and data pre-processing procedures are described which maximise the stability of the algorithm. Examples of applications to both simulated and measured data are presented. The effects of signal bandwidth and measurement noise are discussed as are the improvements in reconstructions over those which are based on simple echo-location ideas. Applications, with examples, of the technique to branched transmission line networks are also described. The time domain method is discussed in relation to the Gelfand-Levitan technique and the inverse normal mode problem.

Chapter 4 serves as an introduction to chapter 5 and reviews cardiac electrophysiology, cardiac arrhythmias and the clinical measurements which are used to assist in arrhythmia diagnosis.

Cardiac arrhythmias are sometimes life threatening and are often difficult to diagnose. The inverse problem in cardiac arrhythmia diagnosis is to determine the mechanism and physiological basis of an arrhythmia from measurements of the surface and intracardiac electrocardiograms under various conditions. Because cardiac rhythm processes are so complex and the amount of clinical data measured is relatively small, it is often not possible for a cardiologist to find a unique mechanism for an arrhythmia. When a system is complex and the data obtained is sparse, inverse problems can usually only be usefully solved by model fitting procedures. Modelling provides a convenient way of including the available *a priori* knowledge in the solution. Chapter 5 describes a modelling approach to arrhythmia diagnosis. The important parameters describing the conduction of electrical

impulses (which control heart rhythm) in the cardiac conduction system are identified. These are used as the basis of an interactive computer model of the conduction system which a cardiologist can use to assist in arrhythmia diagnosis. Means of estimating the model parameters from clinical data are described. Examples of the use of the model to simulate patients' arrhythmias are presented.

The thesis concludes with chapter 6, which presents conclusions and suggestions for future research.

During the course of the work reported in this thesis the following papers and presentations have been prepared.

R.H.T. Bates, G.C. McKinnon and R.P. Millane, Conformations of DNA Compatible with Available Diffraction Data, Presented at the Sci. Meeting of the Christchurch Med. Res. Soc., Christchurch, N.Z., July 1978. Abstract: N.Z. Med. J. (1979), 89, 190.

R.H.T. Bates, G.C. McKinnon and R.P. Millane, A New Look at B-DNA Diffraction Data, Research Report, Electrical Engineering Dept., University of Canterbury, N.Z., Nov. 1978.

R.H.T. Bates, G.C. McKinnon, R.P. Millane and G.A. Rodley, Revised Interpretations of the Available X-ray Data for B-DNA, Pramāna (1980), 14, 233-252.

R.P. Millane and P.J. Bones, A Computer Model of the Cardiac Specialised Conduction System, Presented at the 20th Conf. on Phys. Sci. and Eng. in Med. and Biol., Christchurch, N.Z., Aug. 1980. Abstract: Conf. Proc., p. 37.

R.P. Millane, P.J. Bones, H. Ikram and R.H.T. Bates, Use of a Computer Model to Assist in the Diagnosis of Cardiac Arrhythmias, Presented at the Ann. Sci. Meeting of the Cardiac Soc. of Australia and New Zealand, Dunedin, N.Z., Sept. 1980. Abstract: N.Z. Med. J. (1980), 92, 404.

R.P. Millane, P.J. Bones, H. Ikram and R.H.T. Bates, A Computer Model of Cardiac Conduction, Australasian Phys. Eng. Sci. Med. (1980), 3, 205-209.

R.H.T. Bates and R.P. Millane, Time Domain Approach to Inverse Scattering, IEEE Trans. Ant. Prop. (1981), AP-29, 359-363.

R.P. Millane and G.A. Rodley, Stereochemical Details of the Side-by-Side Model for DNA, Nucleic Acids Res. (1981), 9, 1765-1773.

R.P. Millane and R.H.T. Bates, Inverse Methods for Branched Ducts and Transmission Lines, Proc. IEE on Communications, Radar and Signal Processing (1981), in press.

R.P. Millane, P.J. Bones, M. Flewellen, H. Ikram and R.H.T. Bates, Possible Mechanisms for two Variants of the Wolff-Parkinson-White Syndrome Suggested by Computer Modelling, Presented at the Sci. Meeting of the Christchurch Med. Res. Soc., Christchurch, N.Z., Nov. 1981.

R.P. Millane, G.A. Rodley and G.F. Rodley, Refinement of the Side-by-Side Model for DNA, J. Roy. Soc. N.Z. (1982), in press.

G.A. Rodley, R.P. Millane, G.C. McKinnon and R.H.T. Bates, Use of Axial Pattersons in Assessment of Compatibility of Alternative B-DNA Structures with Fibre X-ray Data, submitted to J. Mol. Biol. (1981).

J.H.T. Bates, W.R. Fright, R.P. Millane, A.D. Seagar, A.E. McKinnon and R.H.T. Bates, Subtractive Image Restoration III: Some Practical Applications, to be submitted to Optik (1982).

R.P. Millane and P.J. Bones, A Numerical Model for Simulating Arrhythmias, to be submitted to PACE (1982).

TABLE OF CONTENTS

1. INVERSE SCATTERING

1.1 INTRODUCTION

In many situations of scientific and technical importance, direct observation of a physical system is either impossible or impracticable. In such cases one often wishes to determine the physical properties of a system from a remote location. Problems of this sort are variously referred to as imaging, remote probing or non-invasive measurement, and they all lie within the realm of inverse problems. The term "inverse problems" is used to distinguish them from direct problems. If one knows the properties of a system and the stimulus applied to it, the direct problem consists of determining the system's response. However in a typical inverse problem, the response of the system to a known stimulus is measured and as many as possible of the system properties are to be determined.

Inverse problems are generally more difficult mathematically, computationally and conceptually than the corresponding direct problems. One of the main reasons for the conceptual difficulty is that they reverse our classical notion of cause and effect. Many inverse problems belong to the class of improperly-posed problems (Deschamps and Cabayan, 1972) in which the solution may depend uniquely but not continuously on the data. Small errors in the data can lead to large errors in the solution unless particular care is taken. Experimental data inevitably contains errors and noise, so this feature of inverse problems is important in practical situations.

In many imaging applications, wave-like radiation is used to probe the unknown system. By measuring the resulting radiation field, one hopes to determine some of the properties of the system. Obviously the interaction between the system and the radiation must be a function of these properties in order for them to be imaged. Inverse problems of this type, involving measurement of the scattering (diffraction) of radiation by material bodies (scatterers) to determine some of their physical properties, are known as inverse scattering problems. Inverse source problems are closely related to inverse scattering problems. Inverse source problems involve the determination of the radiation source given

the properties of the scatterer and measurements of the directly radiated and scattered fields. In many practical inverse source problems the source is located inside the scatterer. Inverse scattering and inverse source problems are often referred to as active and passive sounding respectively (see Colin, 1972). In some cases both the properties of the scatterer and the source need to be determined from the measurements and this adds another order of difficulty to the problem. An example of this is the determination of the structure of the earth from seismic recordings made at the surface, which requires that at least some of the characteristics of the initially unknown earthquake source also be inferred.

Following Bates (1977), the generalised inverse scattering problem may be described in the following way. Consider a space S partitioned into two regions T and Ω (see Fig. 1.1) such that T is exterior to Ω and $S = T \cup \Omega$ where $\cup$ denotes set union. The physical properties of the medium in T are assumed to be known completely and for simplicity it is assumed that T contains free space. A set of functions χ describes the inhomogeneities in the scattering region Ω. An incident wave field Ψ_o is introduced into S and the scattered field Ψ_s is measured at a set of points in the measurement region T. The scattered field is defined by

$$\Psi = \Psi_o + \Psi_s \tag{1.1}$$

where Ψ is called the total field. Note that Ψ reduces to Ψ_o when the medium in Ω is free space, i.e. there are no scatterers. The scattered field contains all the observable information about χ.

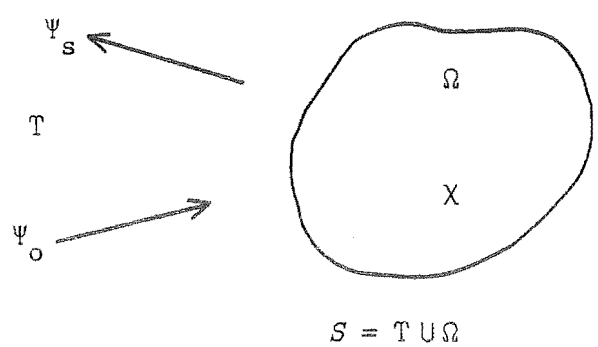


Fig. 1.1 Partitioning of the space S for the generalised inverse scattering problem.

The direct scattering problem consists of determining Ψ_s throughout S given Ψ_o and χ. This operation may be expressed as

$$\Psi_s = \Lambda(\Psi_o, \chi) \tag{1.2}$$

where the direct scattering operator Λ represents a set of mathematical operations on Ψ_o and χ. The operator Λ describes the interaction between the radiation and the scatterers. The inverse scattering problem is to determine χ given Ψ_o and Ψ_s in T, which may be expressed as

$$\chi = \tilde{\Lambda}(\Psi_o, \Psi_s) \tag{1.3}$$

where $\tilde{\Lambda}$ denotes the inverse scattering operator. Equations (1.2) and (1.3) illustrate why inverse scattering problems are more difficult than direct problems, because both χ and Ψ_s have to be inferred within Ω.

In most practical situations the set of functions χ must satisfy a number of constraints in order that the scatterers be physically reasonable and are consistent with other independent information. These constraints form the available *a priori knowledge* of χ. The *a priori* knowledge plays an essential role in inverse problems by reducing the number of possible solutions χ. In fact, in most practical situations, if no *a priori* information is available, the solution would be so non-unique as to be useless. In dealing with inverse problems the mathematical questions of existence, uniqueness and stability are of prime importance.

There are two general approaches to inverse scattering problems. Firstly, $\tilde{\Lambda}$ (or an approximation to $\tilde{\Lambda}$) may be constructed and (1.3) used to determine χ. This procedure is what is usually meant by "inverse scattering". Secondly, χ may be modelled as a set of convenient functions with free parameters. These parameters are fitted to the measurements by repeated solution of the direct problem to check the model against the data until they become consistent. In a strict sense methods of this second sort are model fitting procedures rather than inverse scattering techniques. However in most practical situations, solution methods fall somewhere in between these two extremes, depending on the complexity of the problem and the amount of *a priori* information available.

At the present time, most practical methods of solving inverse scattering problems are based on echo location ideas. A pulse type signal is transmitted and the echoes scattered back from the object are received. The position of a scatterer is estimated from the time between emitted and received pulses and the directional characteristics of the transmitting

and receiving transducers. The "sizes" (or more precisely the "cross-sections") of the scatterers are estimated from the amplitudes of the received pulses. Examples of applications based on these ideas are sonar, radar, seismology and echoradiography. However these techniques assume that the field propagates along straight lines at a known constant speed. If the field passes through a continuously variable medium then ray curvature, variable propagation speed and diffraction effects may become significant. Inverse scattering theory involves examining the possibility of using more rigorous approaches to the problem. Key comprehensive references to inverse scattering are Colin (1972), Chadan and Sabatier (1977), Baltes (1978), Weston (1978), Baltes (1980) and Boerner *et al.* (1981).

In this chapter a number of inverse scattering techniques are reviewed. In §1.2 and §1.3 the wave equations describing electromagnetic and acoustic direct scattering are derived. Section 1.4 deals with propagation in ducts and transmission lines. The rest of the chapter deals with the inverse problem. Sections 1.5 and 1.6 describe two important approximate approaches to inverse scattering, and §1.7 deals with the application of one of these to X-ray crystallography. Aspects of three so-called "exact" solutions to the inverse scattering problem are discussed in §1.8 to §1.10. In §1.11 one of the most important model fitting type procedures is described. The relationship between the techniques described in this chapter and those discussed in chapters 2 and 3 is summarised in §1.12.

It is worth pointing out that techniques considered in this thesis are applicable to penetrable scatterers. The problem of the determination of the location and shape of inpenetrable or perfectly reflecting bodies is just as important and the reader is referred to Colin (1972) and Boerner *et al.* (1981) for references to this topic.

1.2 ELECTROMAGNETIC SCATTERING

The electromagnetic (EM) field is described by Maxwell's equations (Jones, 1964, ch. 1)

$$\left.\begin{aligned} \nabla \times \underline{E} &= -\dot{\underline{B}} \\ \nabla \times \underline{H} &= \dot{\underline{D}} + \underline{J} \\ \nabla \cdot \underline{D} &= \rho \end{aligned}\right\} \quad (1.4)$$

and

$$\nabla \cdot \underline{B} = 0$$

where the field quantities $\underline{E}$, $\underline{H}$, $\underline{D}$, $\underline{B}$, $\underline{J}$ and ρ are the electric field intensity, magnetic field intensity, electric flux density, magnetic flux density, electric current density and charge density respectively. They are functions of position denoted by the vector $\underline{x}$ and time t. A dot above a quantity denotes differentiation with respect to time. The constitutive relations

$$\left.\begin{aligned} \underline{B} &= \mu\underline{H} \\ \underline{D} &= \varepsilon\underline{E} \\ \text{and} \\ \underline{J} &= \sigma\underline{E} \end{aligned}\right\} \quad (1.5)$$

describe the behaviour of <u>isotropic linear</u> matter in the presence of the EM field where μ, ε and σ are the permeability, permittivity and conductivity respectively. Conservation of charge leads to the continuity equation

$$\nabla \cdot \underline{J} + \dot{\rho} = 0 \ . \quad (1.6)$$

Assuming that the medium is source free (i.e. $\rho = \underline{J} = 0$, which implies $\sigma = 0$) and time-invariant ($\dot{\varepsilon} = \dot{\mu} = 0$) then, making use of (1.5), (1.4) reduces to

$$\nabla \times \underline{E} = -\mu\dot{\underline{H}} \quad (1.7)$$

$$\nabla \times \underline{H} = \varepsilon\dot{\underline{E}} \quad (1.8)$$

$$\nabla\cdot(\varepsilon\underline{E}) = 0 \quad (1.9)$$

and

$$\nabla\cdot(\mu\underline{H}) = 0 \quad (1.10)$$

Taking the curl of (1.7) and using the vector identity

$$\nabla \times \nabla \times \underline{F} \equiv \nabla(\nabla \cdot \underline{F}) - \nabla^2\underline{F} \quad (1.11)$$

gives

$$\nabla^2\underline{E} - \nabla(\nabla \cdot \underline{E}) - \nabla\times(\mu\dot{\underline{H}}) = 0 \ . \quad (1.12)$$

Expanding (1.9) and substituting into (1.11) leads to

$$\nabla^2\underline{E} + \nabla(\underline{E} \cdot \nabla \ln\varepsilon) - \mu\nabla \times \dot{\underline{H}} - \nabla\mu \times \dot{\underline{H}} = 0 \ . \quad (1.13)$$

Substituting from (1.7) and (1.8) allows (1.13) to be reduced to

$$\nabla^2\underline{E} - \mu\varepsilon\ddot{\underline{E}} + \nabla(\underline{E} \cdot \nabla \ln\varepsilon) + (\nabla \ln\mu) \times (\nabla \times \underline{E}) = 0 \ . \quad (1.14)$$

The differential equation for $\underline{H}$ is found by similar reasoning to be

$$\nabla^2\underline{H} - \mu\varepsilon\ddot{\underline{H}} + \nabla(\underline{H}\cdot\nabla\ln\mu) + (\nabla\ln\varepsilon)\times(\nabla\times\underline{H}) = 0 \tag{1.15}$$

It is convenient to define the relative permittivity ε_r by

$$\varepsilon = \varepsilon_r\varepsilon_o \tag{1.16}$$

where ε_o is the permittivity of free space. Most materials are non-magnetic ($\mu = \mu_o$ = permeability of free space) so that (1.14) becomes

$$\nabla^2\underline{E} - (\nu/c)^2\ddot{\underline{E}} + \nabla(\underline{E}\cdot\nabla\ln\varepsilon) = 0 \tag{1.17}$$

where the refractive index ν and free space velocity c are defined by

$$\nu = \varepsilon_r^{\frac{1}{2}} \tag{1.18}$$

and

$$c = (\varepsilon_o\mu_o)^{-\frac{1}{2}} . \tag{1.19}$$

When ε is a function of only one dimension then a scalar wavefunction $\Psi = \Psi(\underline{x},t)$ can be chosen which is the component of $\underline{E}$ perpendicular to this dimension and (1.17) reduces to the scalar wave equation

$$\nabla^2\Psi - (\nu/c)^2\ddot{\Psi} = 0 . \tag{1.20}$$

Furthermore, if ε varies arbitrarily in three dimensions but the spatial rate of change of ε is small enough compared to that of $\underline{E}$, then the last term in (1.17) may be neglected so that it reduces to

$$\nabla^2\underline{E} - (\nu/c)^2\ddot{\underline{E}} = 0 . \tag{1.21}$$

Note that the vector components of $\underline{E}$ in (1.21) are uncoupled so that it can be split into three scalar wave equations.

Equation (1.20) is called the time domain or time dependent form of the wave equation. If the field quantities are time harmonic with time dependence $\exp(i\omega t)$, where ω is the angular frequency, then (1.20) becomes

$$\nabla^2\psi + k^2\nu^2\psi = 0 \tag{1.22}$$

where k is the free space wavenumber defined by

$$k = \omega/c \tag{1.23}$$

and $\psi = \psi(\underline{x},k)$ is the temporal Fourier transform (Bracewell, 1978) of $\Psi(\underline{x},t)$ given by

$$\psi(\underline{x},k) = \int_{-\infty}^{\infty} \Psi(\underline{x},t) \exp(i2\pi kct)\, dt \ . \tag{1.24}$$

Equation (1.22) describes the propagation of EM waves under the same restrictions as described above. Equation (1.22) is often called the reduced wave equation or the Helmholtz equation.

1.3 ACOUSTIC SCATTERING

Consider the displacement $s = s(x)$ of an element of an elastic medium from its mean position. The medium is assumed to be a perfect fluid so that it cannot support shear stresses. The strains developed in the medium are assumed to be small so that it is linear and strain is proportional to stress (Hooke's law). Hooke's law may be written as

$$P = -\kappa \nabla \cdot \underline{s} \tag{1.25}$$

where $P = P(\underline{x},t)$ is the excess pressure and κ is called the bulk modulus of elasticity. The force on this element is given by

$$-\iint_A (P\hat{\underline{n}})\, dA = -\iiint_V \nabla P\, dV \tag{1.26}$$

where A and V are the surface area and volume of the element respectively and $\underline{n}$ is the outward normal to A. Taking the limit as $V \sim 0$ shows that the force on the element is equal to $-\nabla P$. Hence Newton's third law for the element is

$$-\nabla P = \rho \ddot{\underline{s}} \tag{1.27}$$

where ρ is the density. Combining (1.25) and (1.27) gives

$$\kappa \nabla(\nabla \cdot \underline{s}) = \rho \ddot{\underline{s}} \ . \tag{1.28}$$

Since the medium cannot support shear stresses, the vector $\underline{s}$ must be irrotational (i.e. $\nabla \times \underline{s} = 0$) and so may be written as the gradient of a scalar Ψ, i.e.

$$\underline{s} = \nabla \Psi \tag{1.29}$$

where Ψ is called the velocity potential. Substituting from (1.29) into (1.28) gives

$$\nabla[\nabla^2\Psi - (1/c^2)\ddot{\Psi}] = 0 \tag{1.30}$$

where the propagation speed c is given by

$$c = (\kappa/\rho)^{\frac{1}{2}} . \tag{1.31}$$

Since the time invariant part of Ψ is of no interest, (1.30) can be integrated to give the wave equation

$$\nabla^2\Psi - (1/c^2)\ddot{\Psi} = 0 . \tag{1.32}$$

Taking the divergence of (1.27) and substituting for $\nabla\cdot\underline{s}$ from (1.25) gives

$$\nabla^2 P - (1/c^2)\ddot{P} = 0 , \tag{1.33}$$

since ρ is approximately constant for small amplitude disturbances, so the excess pressure satisfies the same wave equation. Inspection of (1.21) and (1.32) shows that, under the appropriate conditions, small amplitude acoustic waves and EM waves satisfy wave equations of the same form.

Equation (1.32) describes acoustic waves which propagate in a gas or fluid. They are scalar waves and are called pressure waves or p-waves. Elastic solids, however, support shear stresses and hence shear waves or s-waves propagate in addition to p-waves. Seismic waves which propagate in the earth consist of both p-waves and s-waves (Bullen, 1963). The scattering of s-waves is described by a vector wave equation and they travel slower than the corresponding p-waves.

1.4 WAVES IN DUCTS AND TRANSMISSION LINES

1.4.1 Electromagnetic Waves on Transmission Lines

The propagation of EM waves on transmission lines (for example a line consisting of two separated conductors) may be described in terms of distributed circuit parameters (Jordan and Balmain, 1968, ch. 7). The transverse EM (TEM) mode that propagates in waveguides which consist of two separate conducting structures (for example co-axial lines consisting of an inner and outer conductor, but not hollow cylindrical waveguides

consisting of one conductor) may also be described using non-dispersive transmission line theory. The TEM mode is the only mode which propagates if the highest frequency present is less than the cutoff frequency of the lowest order waveguide mode (Jordan and Balmain, 1968, ch. 7). Transmission line modes are non-dispersive whereas waveguide modes are dispersive (refer to §1.8.1 for a discussion of dispersion).

An EM wave propagating on a transmission line may be characterised by the voltage $V = V(x,t)$ between the two conductors or the current $I = I(x,t)$ flowing in one of the conductors (necessarily equal to the negative of the current flowing in the other conductor) where the coordinate x denotes position on the line. A non-uniform line is considered, which means that the cross-sectional geometry and the medium in which it is embedded may vary arbitrarily with x. It is assumed that the line is lossless which means that the resistivity of the conductors and the conductivity of the medium in which they are embedded are zero. Hence the line may be characterised by a distributed capacitance and inductance per unit length, denoted by $C = C(x)$ and $L = L(x)$ respectively. The voltage and current on the line satisfy the telegraphist's equations (Jordan and Balmain, 1968, ch. 7)

$$\partial V/\partial x = -L\dot{I} \tag{1.34}$$

and

$$\partial I/\partial x = -C\dot{V} \ . \tag{1.35}$$

Taking partial derivatives of (1.34) and (1.35) with respect to x and t respectively allows I to be eliminated, giving

$$\partial^2 V/\partial x^2 - (d(\ln L)/dx)\ \partial V/\partial x - LC\ddot{V} = 0 \ . \tag{1.36}$$

It is convenient to define the characteristic impedance $\zeta = \zeta(x)$ and refractive index $\nu = \nu(x)$ as functions of position on the line by

$$\zeta = (L/C)^{\frac{1}{2}} \tag{1.37}$$

and

$$\nu = c(LC)^{\frac{1}{2}} \tag{1.38}$$

where the constant c is the velocity of propagation when the wires are of uniform spacing and cross-section and are embedded in free space. Equations (1.37) and (1.38) allow (1.36) to be written as

$$\partial^2 V/\partial x^2 - (dG/dx)\ \partial V/\partial x - (\nu/c)^2\ddot{V} = 0 \tag{1.39}$$

where

$$G = G(x) = \ln(\zeta\nu/c) \ . \tag{1.40}$$

1.4.2 Sound Waves in Ducts

Consider the propagation of sound waves in a duct or tube filled with a homogeneous fluid and whose cross-sectional shape varies with position x along the axis of the duct or tube. The wave propagating in the duct must be a plane wave if it is to be a function of only one spatial dimension. If the cross-sectional dimensions of the duct do not change too rapidly with x and are small compared to the shortest wavelength of the wave then the wave motion is approximately planar (Morse, 1948, §24). The propagation can then be accurately described by transmission line theory.

Consider a thin shell of fluid in the duct. Let $s = s(x)$ be the displacement (which is in the x direction since the wave is planar) of this shell and let $A = A(x)$ be the cross-sectional area of the duct. Hence for this elemental shell (1.25) becomes

$$P = -(\kappa/A)\ \partial(As)/\partial x \tag{1.41}$$

and (1.27) becomes

$$-\partial P/\partial x = \rho\ddot{s} \ . \tag{1.42}$$

Differentiating (1.42) with respect to x and substituting for $\partial s/\partial x$ by expanding (1.41) gives

$$\partial^2 P/\partial x^2 + (dH/dx)\ \partial P/\partial x - (1/c^2)\ddot{P} = 0 \tag{1.43}$$

where

$$H = H(x) = \ln A \tag{1.44}$$

and c is the free space propagation speed defined by (1.31). The characteristic impedance of the duct is defined by

$$\zeta = c/A \ . \tag{1.45}$$

Inspection of (1.39), (1.40), (1.43) and (1.44) shows the equivalence between the EM transmission line and the acoustic duct. The refractive index of the duct is constant (equal to unity because of the way c has been defined) because the duct is filled with a homogeneous fluid.

1.5 THE BORN INVERSE SCATTERING APPROXIMATION

The Born approximation to inverse scattering, sometimes called the inverse Born approximation, is based on the Born (or Rayleigh-Gans) approximate method of solving the direct scattering problem (Jones, 1964, §6.13). It is the basis of X-ray crystallography which is discussed in §1.7. The Born approximation is a frequency domain or spectral technique as it is formulated for the scattering of monochromatic waves described by (1.22), which is repeated here;

$$\nabla^2\psi + k^2\nu^2\psi = 0 \ . \tag{1.46}$$

Equation (1.46) may be written in the form

$$\nabla^2\psi + k^2\psi = -S = -k^2(\nu^2 - 1)\psi \tag{1.47}$$

where $S = S(\underline{x})$ represents equivalent sources which have been called polarization sources by Bates and Ng (1972) since they represent the polarization of the medium by the total field. Equation (1.47) may be solved for ψ using the Green's function technique (Morse and Feshbach, 1953, ch. 7) so that

$$\psi(\underline{x},k) = \psi_o(\underline{x},k) + \iiint_\Omega S(\underline{\xi})\ G(\underline{x},\underline{\xi})\ d\underline{\xi} \tag{1.48}$$

where ψ_o is a solution of the homogeneous or free space equation

$$\nabla^2\psi_o + k^2\psi_o = 0 \tag{1.49}$$

and Ω is defined in §1.1. Reference to (1.1) and (1.49) shows that ψ_o is the incident field defined in §1.1. The free space Green's function, G, for (1.47) is given by (Cowley, 1975, §1.5)

$$G(\underline{x},\underline{\xi}) = \exp(ik|\underline{x}-\underline{\xi}|)/(4\pi|\underline{x}-\underline{\xi}|) \ . \tag{1.50}$$

Substituting from (1.47) and (1.50) into (1.48) and making use of (1.1) and (1.49) allows the scattered field ψ_s to be written as

$$\psi_s(\underline{x},k) = \frac{k^2}{4\pi}\iiint_\Omega [\nu^2(\underline{\xi})-1]\ [\psi_o(\underline{\xi},k) + \psi_s(\underline{\xi},k)]\ \frac{\exp(ik|\underline{x}-\underline{\xi}|)}{|\underline{x}-\underline{\xi}|}\ d\underline{\xi} \ . \tag{1.51}$$

The direct problem can be solved using (1.51) which is an integral equation for ψ_s, since ψ_o and ν are known. However the difficulty with

using (1.51) to solve the inverse problem for ν is that ψ_s is known only in T but not in Ω. The Born approximation (more strictly, the first term of the Born series - see Cowley, 1975, §1.5) consists of assuming that the scattering is so weak that

$$\psi_s << \psi_o \quad \text{for} \quad \underline{x} \in S \tag{1.52}$$

and (1.51) may be approximated by

$$\psi_s(\underline{x},k) = \frac{k^2}{4\pi} \iiint_{\Omega} [\nu^2(\underline{\xi}) - 1]\ \psi_o(\underline{\xi},k)\ \frac{\exp(ik|\underline{x} - \underline{\xi}|)}{|\underline{x} - \underline{\xi}|}\ d\underline{\xi} \ . \tag{1.53}$$

Equation (1.53) may be solved for ν since ψ_s is only required in the measurement region T.

A particularly simple form of (1.53) results if the incident field is a plane wave

$$\psi_o(\underline{x},k) = \exp(ik\underline{\hat{k}}\cdot\underline{x}) \tag{1.54}$$

where the vector wavenumber $\underline{k}$ is defined by $\underline{k} = k\underline{\hat{k}}$, $\hat{}$ denotes a unit vector and $\underline{\hat{k}}$ is the direction of propagation of the plane wave. Assuming that ψ_s is measured in the far field ($|\underline{x}| >> |\underline{\xi}|$) then (1.53) may be solved using Fourier transforms. Substituting from (1.54) and applying (1.53) in the far field gives

$$\psi_s(\underline{x},k) = \frac{k^2 \exp(ik|\underline{x}|)}{4\pi|\underline{x}|} \iiint_{\Omega} [\nu^2(\underline{\xi}) - 1]\ \exp[i(\underline{k} - k\underline{\hat{x}})\cdot\underline{\xi}]\ d\underline{\xi} \ . \tag{1.55}$$

If ψ_s is measured at a constant radius from the origin ($|\underline{x}|$ is constant) and the observation coordinates are transformed from $\underline{x}$ to $\underline{u}$ (sometimes called the scattering vector, see Fig. 1.2) defined by

$$\underline{u} = (\underline{k} - k\underline{\hat{x}})/2\pi \ , \tag{1.56}$$

then (1.55) may be written as

$$F(\underline{u}) = \iiint_{-\infty}^{\infty} [\nu^2(\underline{\xi}) - 1]\ \exp(i2\pi\underline{u}\cdot\underline{\xi})\ d\underline{\xi} \ . \tag{1.57}$$

$F = F(\underline{u})$ is equal to ψ_s divided by the scale factor outside the integral in (1.55) and the region of integration has been changed to all of space because $\nu^2 - 1 = 0$ outside Ω. Equation (1.57) is a Fourier transform and hence can be inverted to give ν explicitly in terms of F:

$$\nu^2(\underline{x}) - 1 = \iiint_{-\infty}^{\infty} F(\underline{u})\ \exp(-i2\pi\underline{x}\cdot\underline{u})\ d\underline{u} \ . \tag{1.58}$$

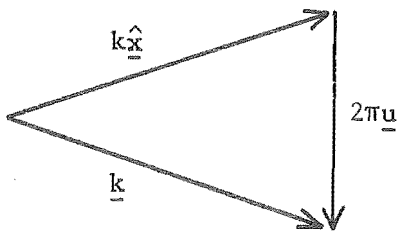


Fig. 1.2 Definition of the scattering vector $\underline{u}$.

For the inverse Born approximation to be useful, (1.55) must hold. This means firstly, that $(\nu-1)$ must be small so that the scattering is weak and the amplitude of ψ_s is small in comparison with that of ψ_o. Secondly, $(\nu-1)L$, where L is the largest linear dimension of Ω, must be small to ensure that the additional phase accumulated by ψ compared to ψ_o is small. This second restriction is important if L is greater than a few wavelengths and can cause large errors in the reconstruction of ν if it is violated (Vezzetti and Aks, 1979). Hence the inverse Born approximation can be expected to be useful only if the scatterer is either very tenuous or consists of very small scatterers.

Vezzetti and Aks (1979) describe the following iterative scheme to improve reconstructions based on the inverse Born approximation. The reconstructed ν calculated using (1.58) is used to estimate the average refractive index, $\langle\nu\rangle$, in Ω. In order to help account for the additional phase shift of ψ compared to ψ_o, the total field on the RHS of (1.51) is replaced by $\exp(i\langle\nu\rangle\underline{k}\cdot\underline{x})$ rather than ψ_o. The FT relationship (1.58) still applies and a new reconstructed ν is calculated and the process repeated until ν settles down. The results of Vezzetti and Aks (1979) indicate that, if the initial reconstruction is reasonably accurate, this procedure may give useful improvements in the reconstruction. However iterative correction procedures of this type must be treated with caution as convergence is not assured. It sometimes happens that even if the procedure converges, the corrected reconstruction is less accurate than that initially obtained.

1.6 THE RYTOV INVERSE SCATTERING APPROXIMATION

The Rytov approximation to inverse scattering is similar to the inverse Born approximation as it is based on a weak scattering solution to the direct problem. The difference between these two approximations stems from defining Rytov's scattered $\phi_s = \phi_s(\underline{x},k)$ by

$$\phi_s = \psi_o(\ln\psi - \ln\psi_o) \tag{1.59}$$

rather than the conventional scattered field $\psi_s = \psi - \psi_o$. In order to derive the wave equation satisfied by ϕ_s, the functions γ, γ_o and γ_s are defined by

$$\left.\begin{aligned} \psi &= \exp(\gamma) \\ \text{and} \quad \psi_o &= \exp(\gamma_o) \\ \gamma_s &= \gamma - \gamma_o \end{aligned}\right\} \tag{1.60}$$

so that

$$\phi_s = \psi_o\gamma_s \ . \tag{1.61}$$

Substituting for ψ_o from (1.60) into (1.49) shows that γ_o satisfies

$$\nabla^2\gamma_o + \nabla\gamma_o \cdot \nabla\gamma_o + k^2 = 0 \ . \tag{1.62}$$

Substituting for ψ from (1.60) into (1.46) and making use of (1.62) gives

$$\nabla^2\gamma_s + 2\nabla\gamma_o \cdot \nabla\gamma_s + \nabla\gamma_s \cdot \nabla\gamma_s + k^2(\nu^2 - 1) = 0 \ . \tag{1.63}$$

The Rytov approximation is predicated on the scattering being weak, so that

$$\nabla\gamma_s << \nabla\gamma_o \tag{1.64}$$

which means that (1.63) reduces to

$$\nabla^2\gamma_s + 2\nabla\gamma_o \cdot \nabla\gamma_s + k^2(\nu^2 - 1) = 0 \ . \tag{1.65}$$

Making use of (1.61), (1.49), (1.60) and (1.65) shows that Rytov's scattered field satisfies

$$\nabla^2\phi_s + k^2\phi_s = -k^2(\nu^2 - 1)\ \psi_o \ . \tag{1.66}$$

Since the term on the RHS of (1.66) depends only on the incident field, (1.66) may be transformed to an integral equation for $(\nu^2 - 1)$ similar to (1.53). The integral equation may be converted to a FT under the conditions described in §1.5. It is claimed (Chernov, 1960, §16) that the Rytov approximation is more accurate for extended scatterers than the Born approximation because it only requires that the relative amplitude and additional phase changes of ψ per wavelength must be small. However, as shown by Keller (1969), the Rytov approximation is only superior to the Born approximation when the total field closely approximates a single plane wave. This means that it is only useful when most of the scattering is in the forward direction, or that "refraction predominates over reflection" (Bates *et al.*, 1976).

Bates *et al.* (1976) describe an extension to the Rytov approximation by taking partial account of the term $\nabla\gamma_s \cdot \nabla\gamma_s$ neglected in (1.65). They show that this extension can give increased accuracy for some forward scattering computations and that the inverse problem can still be formulated as an integral equation of the same form as (1.53).

1.7 X-RAY CRYSTALLOGRAPHY

1.7.1 X-rays and Molecular Structure

This section is concerned with the theory of the diffraction of X-rays by crystals, which is the basis of molecular structure determination. It serves as an introduction to the techniques used in chapter 2 to study the structure of DNA. The minimum distance between atoms in a crystal is about 1.5 Å (1 Å = 10^{-10} m). X-rays used in crystallography have a wavelength of about 1.5 Å and so are suitable for imaging the molecular structure of crystals. Technological constraints (the short wavelength of X-rays would demand incredibly small tolerances) have precluded the construction of an "X-ray microscope" which would allow direct imaging of molecules at atomic resolution. Hence when a crystal is irradiated with X-rays the only measurable quantity is the diffraction pattern formed by the scattered X-rays. The information contained in the diffraction pattern has to be processed numerically in order to produce an image.

X-rays are EM radiation which cause charged particles to oscillate. The accelerated particles form a time varying current which re-radiates an EM field. If the frequency of the incident X-rays is much greater than

the resonant frequency of the particles then the scattering is elastic (the scattered and incident X-rays have the same frequency) and is described by Thompson scattering theory (Cowley, 1975, §4.1). Since the scattered amplitude is inversely proportional to the mass of the particle, scattering by the electrons in atoms is much more significant than by the protons. The electron density $f = f(\underline{x})$ acts like a potential and hence satisfies a wave equation of the Schrodinger form (see §1.9.1)

$$\nabla^2\psi + k^2\psi = -f\psi \ . \tag{1.67}$$

Because the space occupied by atoms is so small, the scattering is sufficiently weak (most of the incident X-ray beam passes undeflected through the crystal) that the Born approximation (§1.5) applies. Hence the diffraction pattern $F = F(\underline{u})$ is the Fourier transform of the electron density:

$$F(\underline{u}) = \iiint_{-\infty}^{\infty} f(\underline{x}) \exp(i2\pi \underline{u}\cdot\underline{x})\, d\underline{x} \ . \tag{1.68}$$

In crystallographic terminology, the space containing the scattering vector $\underline{u}$ is called reciprocal space.

1.7.2 Diffraction by Crystals

A crystal is a collection of atoms whose centres are arranged in a three-dimensional array or lattice. The latter is defined as a collection of points called lattice points. The geometry of the crystal lattice is defined by the unit cell, which is the smallest parallelepiped which can be constructed with lattice points at its corners. By repeating the unit cell regularly and indefinitely in three dimensions to fill all space, the crystal lattice is generated. The lattice vectors $\underline{a}$, $\underline{b}$ and $\underline{c}$ are defined by the edges of the unit cell. Lattices encountered in this thesis have lattice vectors which are orthogonal and a unit cell of this type is called orthorhombic.

The electron density, f, of a crystal can be written as a convolution (Bracewell, 1978, ch. 3) of the electron density, e, of a single unit cell, with the lattice points. Hence

$$f(x,y,z) = e(x,y,z) \otimes \sum_{h,k,\ell=-\infty}^{\infty} \delta(x-ha)\ \delta(y-kb)\ \delta(z-\ell c) \tag{1.69}$$

where (x,y,z) defines a Cartesian coordinate system, a, b and c are the lengths of the lattice vectors, $\delta(\cdot)$ is the Dirac delta function, and $\otimes$

denotes convolution. On defining a Cartesian coordinate system (u,v,w) in reciprocal space, (1.68) becomes

$$E(u,v,w) = \iiint_{-\infty}^{\infty} e(x,y,z)\ \exp(i2\pi(ux+vy+wz))\ dxdydz \qquad (1.70)$$

where E(u,v,w) is the diffraction pattern of a single unit cell. However, what is measured is the diffraction pattern F(u,v,w) of the whole crystal, which is obtained by Fourier transforming (1.69) and applying the convolution theorem (Bracewell, 1978, ch. 6), which gives

$$F(u,v,w) = E(u,v,w)\ (1/abc) \sum_{h,k,\ell=-\infty}^{\infty} \delta(u-h/a)\ \delta(v-k/b)\ \delta(w-\ell/c). \qquad (1.71)$$

Hence the observed diffraction pattern is equivalent to the diffraction pattern of a single unit cell sampled at the reciprocal lattice points (h/a, k/b, ℓ/c). The complex amplitudes of these samples are called the structure factors

$$F_{hk\ell} = abc\ F(h/a,k/b,\ell/c)\ . \qquad (1.72)$$

Making use of (1.70) to (1.72) gives

$$F_{hk\ell} = \int_0^c \int_0^b \int_0^a f(x,y,z)\ \exp(i2\pi(hx/a + ky/b + \ell z/c))\ dxdydz\ . \qquad (1.73)$$

Equation (1.71) is an expression of Bragg's law (Sherwood, 1976) which states that significant diffracted intensity occurs only at discrete angles where the diffracted waves are in phase. If the unit cell contains N atoms then the structure factors are given by

$$F_{hk\ell} = \sum_{n=1}^{N} \tilde{F}_{n,hk\ell}\ \exp(i2\pi(hx_n/a + ky_n/b + \ell z_n/c)) \qquad (1.74)$$

where (x_n,y_n,z_n) are the coordinates of the centre of the n^{th} atom, and the $\tilde{F}_{n,hk\ell}$ are the structure factors of the n^{th} atom when it is centred at the origin.

It is conventional crystallographic practice to write the reciprocal lattice indices -h, -k, and -ℓ as $\bar{h}$, $\bar{k}$ and $\bar{\ell}$ respectively. Since the electron density is real it follows that

$$F_{hk\ell} = F^{*}_{\bar{h}\bar{k}\bar{\ell}} \qquad (1.75)$$

where * denotes the complex conjugate.

1.7.3 Structure Determination

Structure determination (the basic goal of X-ray crystallography) involves the determination of the positions of the atoms in the unit cell. Since the electron density of an atom is concentrated close to its nucleus, the peaks in the electron density occur at the atomic positions. The peak amplitudes are proportional to the atomic numbers and so identify the types of the atoms. Hence the structure is "solved" if the electron density, f, is determined to sufficient resolution.

Fourier transforming (1.73) gives

$$f(x,y,z) = \sum_{h,k,\ell=-\infty}^{\infty} F_{hk\ell} \exp(i2\pi(hx/a + ky/b + \ell z/c)) \qquad (1.76)$$

and so the electron density is determined from the structure factors. The operation of computing f using (1.76) is known as Fourier synthesis. However technological constraints preclude measurement of the phase of the diffraction pattern, so what is actually measured are the structure intensities $|F_{hk\ell}|^2$. Hence Fourier synthesis can only be used if the phases of the structure factors can be inferred. This is an example of the phase problem which is encountered in many branches of science (see, for example, Bates, 1978). If it was not for the phase problem, the determination of molecular structures would be relatively straightforward. However crystallographers have developed a number of methods of phase determination which are divided into "direct methods" (see Dolfson, 1961) and "indirect methods" (see Sherwood, 1976, ch. 12). These later methods are used for macromolecular structure determination and are described later in this section.

A useful quantity which can be computed directly from the observed structure intensities $|F_{hk\ell}|^2$, is the Patterson function or Patterson map, P, equal to the FT of the $|F_{hk\ell}|^2$ so that

$$P(x,y,z) = \sum_{h,k,\ell=-\infty}^{\infty} |F_{hk\ell}|^2 \exp(i2\pi(hx/a + ky/b + \ell z/c)) \quad . \qquad (1.77)$$

Substituting for $F_{hk\ell}$ from (1.73) gives

$$P(x,y,z) = \sum_{h,k,\ell=-\infty}^{\infty} A(x - ha, y - kb, z - \ell c) \qquad (1.78)$$

where A is the autocorrelation (Bracewell, 1978, ch. 3) of the electron density of a single unit cell given by

$$A(x,y,z) = \iiint_{-\infty}^{\infty} e(\xi,\eta,\zeta)\ e(\xi-x,\eta-y,\zeta-z)\ d\xi d\eta d\zeta \ . \qquad (1.79)$$

Equation (1.78) shows the difference between the Patterson and the autocorrelation, the former being equal to the sum of autocorrelations centred at each lattice point (see Fig. 1.3) and hence it has the same periodicity as the structure. The autocorrelation, and hence the Patterson, is symmetric about the origin. Equations (1.78) and (1.79) show that the Patterson has peaks at positions corresponding to interatomic vectors in the crystal. The amplitude of each peak is equal to the product of the two atomic numbers of the corresponding atoms. The effect of (1.78) is that the autocorrelation contains contributions only from atom pairs within a single unit cell whereas the Patterson also contains those due to atom pairs in adjacent unit cells.

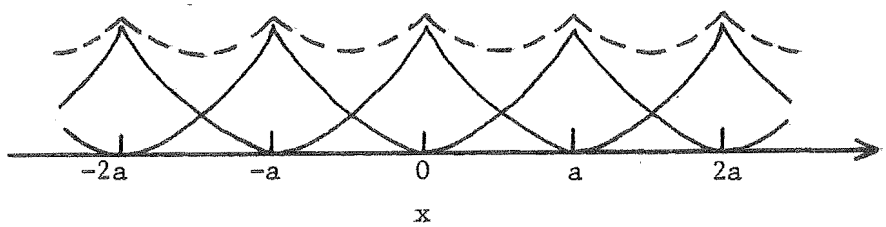


Fig. 1.3 The Patterson (---) and autocorrelations (——) in one dimension.

If the unit cell contains N atoms then the Patterson has N(N-1) peaks. Each Patterson peak is the cross-correlation of two electron density peaks and so the width of a Patterson peak is the sum of the widths of the two corresponding electron density peaks. Hence the degree to which the Patterson may be interpreted to give direct information on atomic positions depends critically on N. If N is very small (i.e. about 10 or so) the Patterson can often be unravelled directly, with the aid of independent chemical knowledge, to determine the atomic positions. However biological molecules may contain over 1000 atoms in the unit cell. The Patterson then contains over 10^6 peaks, many of which overlap, making direct interpretation virtually impossible. However one of the main uses of Patterson maps is phase determination which is discussed next.

The main methods of phase determination, particularly for biological macromolecules, are chemical modification methods (Sherwood, 1976, ch. 12). These all require investigation of the diffraction pattern of at least two isomorphous crystals, one of which must contain a heavy atom (an atom with a large atomic number). Isomorphous crystals have (nearly) identical crystallographic structure and very similar chemical constituents. The diffraction pattern of an isomorphous crystal in which a small number of atoms have been replaced by heavy atoms is measured and the Patterson is computed. Patterson peaks corresponding to interatomic vectors involving the heavy atoms stand out clearly and this usually allows the positions of the heavy atoms to be determined. Let F and F_I be the structure factors, at a point in reciprocal space, of the original sample and the isomorphous sample respectively. Then $F_I = F + F_H$ so that

$$|F_I|^2 = |F|^2 + |F_H|^2 + FF_H^* + F^*F_H \tag{1.80}$$

where F_H is the structure factor of the heavy atoms minus the structure factors of the atoms which they replaced. Equation (1.80) may be written

$$2|F||F_H|\cos(\alpha - \beta) = |F_I|^2 - |F|^2 - |F_H|^2 \tag{1.81}$$

where α and β are the phases of F and F_H respectively. Now $|F|$ and $|F_I|$ are known from the diffraction patterns, and the positions of the heavy atoms are found from the Patterson, allowing $|F_H|$ and β to be computed. Hence (1.81) can be solved for α, the phase of the structure factor F. Unless the crystal has special symmetry properties, three isomorphous samples are required to obtain α unambiguously (Sherwood, 1976, ch. 12).

In some cases it is convenient to choose a heavy atom which has an absorption band close to the X-ray wavelength. The X-rays then suffer anomalous (inelastic) scattering by the heavy atom, so the electron density behaves as a complex quantity and (1.75) is not satisfied. Information from anomalous scattering can assist in phase determination (Ramachandran and Srinivasan, 1970). Once the structure factor phases have been found, the electron density can be computed by Fourier synthesis.

1.7.4 Structure Refinement

The atomic positions obtained as described in §1.7.3 are subject to inaccuracies due to errors in the measured structure intensities, errors in phase determination, truncation of the series in (1.76) because of the

finite number of measured structure factors and uncertainties in inferring the atomic positions from the electron density map. The effect of these errors on the structure can be minimised by refinement procedures. Structure refinement normally proceeds in three stages. Fourier refinement consists of computing structure factor phases from an estimate of the structure and using these, together with the observed magnitudes, to compute a new structure. The process is repeated until there is no significant change in the phases. This is similar to phase correction techniques used in other branches of image processing (see, for example, Fienup, 1978). Following this, a difference Fourier map is computed which is a Fourier synthesis using the difference between the observed and calculated structure amplitudes. Difference Fourier synthesis is used for the more precise location of atomic positions and identification of missing atoms. Finally the atomic positions are adjusted by minimising (using least squares techniques) the crystallographic residual (or R-factor) R, given by

$$R = \left(\sum_m \left| |F_o|_m - |F_c|_m \right| \right) / \sum_m |F_o|_m \qquad (1.82)$$

where $|F_o|_m$ and $|F_c|_m$ are the m^{th} structure amplitudes observed and calculated respectively. At this stage the best possible estimate of the structure should have been obtained.

Often, particularly with biological macromolecules, high quality crystals cannot be obtained, which means that the diffraction data obtained is also of very low quality. In such cases, often the best that can be done is to propose a model for the structure and refine it using least squares refinement.

Structure determination using X-ray crystallography provides a good example of the essential role of *a priori* knowledge mentioned in §1.1. From chemical analysis the number and types of the atoms in the molecule and their electron density functions are known. Stereochemical parameters (geometrical properties such as bond lengths, bond angles and a number of others) must be within fairly narrow tolerances. All this *a priori* information is necessary in order to obtain a reliable solution to the crystallographic inverse scattering problem.

1.8 BATES' SOLUTION TO THE INVERSE SCATTERING PROBLEM

It was shown in §1.2 and §1.3 that, under appropriate conditions, EM and acoustic scattering is accurately described by the Helmholtz equation (1.22). Solutions to the inverse scattering problem described so far have all been approximate as they assume that the scattering is weak. In this section a method proposed by Bates (1975) to exactly solve the inverse scattering problem for the Helmholtz equation is described. This seems to be the only published attempt to solve this problem in more than one spatial dimension. Exact solutions to the inverse scattering problem for the Schrödinger equation are discussed in §1.9.

The method is described for two space dimensions although it is easily extended to three dimensions by invoking spherical Bessel functions and spherical harmonics. A polar coordinate system (r,θ) is introduced in S, and Ω is taken to be a circle of radius a circumscribing the scattering region so that

$$\left.\begin{aligned} \nu &= \nu(r,\theta) \quad , \quad r < a \\ &= 1 \quad , \quad r \geq a \ . \end{aligned}\right\} \tag{1.83}$$

The wavefunction is written as an angular Fourier series

$$\psi(r,\theta,k) = \sum_{m=-\infty}^{\infty} \psi_m(r,k) \exp(im\theta) \ . \tag{1.84}$$

Now ψ can be measured for $r \geq a$ which allows the wave impedance vector $\underline{Z}$ at $r = a$ to be determined. The m^{th} component of $\underline{Z}$ is defined by

$$Z_m(k) = \psi_m(r,k)/\psi_m'(r,k)\Big|_{r=a} \tag{1.85}$$

where the prime denotes the derivative with respect to r. The inverse scattering problem is to determine ν for $r < a$ given $\underline{Z}$. Firstly ν and ψ are written as series of orthogonal functions satisfying the necessary boundary conditions at $r = a$. Since $\nu = \nu(r,\theta)$ is equal to unity at $r = a$, it can be written as a Fourier-Bessel series (Watson, 1966, ch. 18)

$$\nu^2 - 1 = \sum_{n=-\infty}^{\infty} \sum_{p=1}^{\infty} A_{np} J_n(g_{np} r/a) \exp(in\theta) \quad , \quad r < a \tag{1.86}$$

where $J_n(\cdot)$ is the cylindrical Bessel function of the first kind of order n and g_{np} is the p^{th} zero of J_n so that

$$J_n(g_{np}) = 0 \ . \tag{1.87}$$

An appropriate series expansion for ψ which satisfies the impedance boundary condition (1.85) is the Dini series (Watson, 1966, ch. 18)

$$\psi(r,\theta,k) = \sum_{m=-\infty}^{\infty} \sum_{\ell=1}^{\infty} b_{\ell m}(k)\, J_m(h_{\ell m}(k) r/a)\, \exp(im\theta) \tag{1.88}$$

where the $h_{\ell m}(k)$ are defined by

$$J_m(h_{\ell m}) = (h_{\ell m}/a)\, Z_m\, J_m'(h_{\ell m}) \ . \tag{1.89}$$

Note that the $h_{\ell m}$ can be determined, using (1.89), from the measurements Z_m. Using (1.88) and the fact that the J_m satisfy Bessel's differential equation (Watson, 1966, ch. 2), $\nabla^2\psi$ can be written as

$$\nabla^2\psi = - \sum_{m=-\infty}^{\infty} \sum_{\ell=1}^{\infty} b_{\ell m}(h_{\ell m}/a)^2\, J_m(h_{\ell m} r/a)\, \exp(im\theta) \ . \tag{1.90}$$

Substituting from (1.86), (1.88) and (1.90) into (1.22) and using the orthogonality of $\exp(im\theta)$ and $J_m(h_{\ell m} r/a)$ gives

$$\sum_{m=-\infty}^{\infty} \sum_{\ell=1}^{\infty} \left\{ [k^2 - (h_{\ell' m'}/a)^2]\, N_{\ell' m'}\, \delta_{\ell\ell'}\, \delta_{mm'} + k^2 \sum_{p=1}^{\infty} A_{m'-m,p}\, \chi_{\ell,\ell',m,m',m'-m,p} \right\} b_{\ell m} = 0 \tag{1.91}$$

where $\ell' \geq 0$,

$$\chi_{\ell,\ell',m,m',n,p}(k) = \int_0^a J_n(g_{np} r/a)\, J_m(h_{\ell m} r/a)\, J_{m'}(h_{\ell' m'} r/a)\, r\, dr \ , \tag{1.92}$$

$$N_{\ell m}(k) = \int_0^a [J_m(h_{\ell m} r/a)]^2\, r\, dr \tag{1.93}$$

and δ_{mn} is the Kronecker delta defined by

$$\left.\begin{aligned} \delta_{mn} &= 1 \quad , \quad m = n \\ &= 0 \quad , \quad m \neq n \ . \end{aligned}\right\} \tag{1.94}$$

The quantities defined by (1.92) and (1.93) can be computed from the measurements, and (1.91) can be treated as an infinite set of homogeneous, linear, simultaneous equations for the unknown $b_{\ell m}$ with the unknown A_{np} appearing as coefficients. If (1.91) is to have a non-trivial solution then the determinant of the coefficient matrix, which Bates calls the

inverse scattering determinant, must equal zero. The significant property of this determinant is that the only unknowns it contains are the A_{np} which characterise the inhomogeneity ν. If measurements are made at a number, M say, of values of k then M independent determinants are obtained. If M is sufficiently large then the determinants may be truncated to order M and solved simultaneously for M values of the A_{np} thereby giving an estimate of ν. However the determinant is very non-linear in the A_{np} and the practical numerical aspects of the method require further investigation.

At present there is no exact solution to the inverse scattering problem for the HE in more than one dimension. Although the method described above does not provide a closed form solution, it does go some of the way towards formulating such a solution.

1.9 THE GELFAND-LEVITAN TECHNIQUE

1.9.1 The Schrödinger Equation and the Scattering Matrix

The Schrödinger equation (SE)

$$\nabla^2\psi + k^2\psi - V\psi = 0 \tag{1.95}$$

describes the (non-relativistic) quantum mechanical scattering of particles by the potential $V = V(\underline{x})$. The wavefunction ψ characterises the particles and the wavenumber k is related to their energy. In this section the discussion is limited to the one dimensional form of (1.95),

$$\partial^2\psi/\partial x^2 + k^2\psi - V\psi = 0 \ , \tag{1.96}$$

which applies to the scattering of plane waves by a plane stratified medium or, with slight modifications, by a circularly symmetric potential. Equations of the form of (1.96) also describe EM scattering by a plane stratified plasma (Jordan and Ahn, 1979) and monochromatic EM plane wave scattering by a plane stratified dielectric slab where the angle of incidence (rather than the wavenumber) is the spectral variable (Mittra *et al.*, 1972). Fourier transforming (1.96) gives the time domain form

$$\partial^2\Psi/\partial x^2 - (1/c^2)\ddot{\Psi} - V\Psi = 0 \ . \tag{1.97}$$

Note that t in (1.97) is a pseudo-time because in the conventional Schrodinger equation the term $i\partial\Psi/\partial(\text{time})$ occurs rather than the term

$-\partial^2\Psi/\partial(\text{time})^2$.

The difference between the SE (1.96) or (1.97) and the Helmholtz equation (HE) (1.22) or (1.20) is that the function describing the inhomogeneity is multiplied by the wavefunction in the former rather than its second time derivative in the latter. This means that for the Schrödinger equation the propagation is dispersive and furthermore that the wavefront travels at the free space velocity c. Both of these characteristics can be understood by examining the variation of propagation speed with wavenumber. The local propagation speed v inside the scattering region is equal to ω/κ where κ is the local wavenumber. Reference to (1.96) and (1.22) shows that

$$\left.\begin{aligned} v &= c(1 - V/k^2)^{-\frac{1}{2}} \quad \text{for the SE} \\ &= c/\nu \qquad\qquad\quad \text{for the HE .} \end{aligned}\right\} \quad (1.98)$$

Hence the SE is <u>dispersive</u> because v is a function of k, so that, even in a homogeneous medium, the shape of Ψ changes with time as the different frequency components travel at different speeds. However the HE is <u>non-dispersive</u> because v is independent of k and so the wave shape is constant with time in a homogeneous medium. Examination of (1.98) shows, assuming $V \leq 0$, that v increases from 0 to c as $|k|$ varies from 0 to ∞, so the wavefront of Ψ travels at the free space speed c. Hence the depth of penetration of Ψ into the scattering region after a particular time is independent of the inhomogeneity V. However for the HE the wavefront travels at a speed c/ν and so the depth of penetration is a function of the inhomogeneity ν.

The scattering of one-dimensional scalar waves may be described with the aid of the <u>scattering matrix</u> (or <u>S-matrix</u>). Consider the scattering of monochromatic waves by a potential which vanishes at infinity, i.e. $V(x) \sim 0$ as $x \sim \pm\infty$. As $|x|$ approaches infinity, the medium approaches free space and the wavefunction satisfies the free space equation, so that

$$\left.\begin{aligned} \psi &= A(k)\ \exp(ikx) + B(k)\ \exp(-ikx) \quad \text{as} \quad x \sim +\infty \\ &= C(k)\ \exp(ikx) + D(k)\ \exp(-ikx) \quad \text{as} \quad x \sim -\infty. \end{aligned}\right\} \quad (1.99)$$

The scattering matrix $S = S(k)$ transforms the incoming waves into the outgoing waves:

$$\begin{pmatrix} A \\ D \end{pmatrix} = S \begin{pmatrix} C \\ B \end{pmatrix} = \begin{pmatrix} s_{11}(k) & s_{12}(k) \\ s_{21}(k) & s_{22}(k) \end{pmatrix} \begin{pmatrix} C \\ B \end{pmatrix}. \quad (1.100)$$

Applying (1.100) to an incident wave from first the left ($x = -\infty$) and then the right ($x = +\infty$) shows that s_{21} and s_{11} (and s_{12} and s_{22}) are the reflection and transmission coefficients respectively from the left (and right). The principle of reciprocity (Jones, 1964, §1.3.2) ensures that

$$s_{11} = s_{22} \tag{1.101}$$

and energy conservation requires that S be Unitary,

$$(S^*)^T S = I \ , \tag{1.102}$$

where the superscript T denotes matrix transposition and I is the unit matrix. $(S^*)^T$ is often called the Hermitian conjugate or adjoint of S. Carrying out the operations in (1.102) gives

$$\left.\begin{aligned} &|s_{11}|^2 + |s_{21}|^2 = |s_{12}|^2 + |s_{22}|^2 = 1 \\ \text{and} \\ &s_{11}^* s_{12} + s_{21}^* s_{22} = 0 \ . \end{aligned}\right\} \tag{1.103}$$

It can be shown (Kay, 1972) that (1.101) and (1.103), together with physically reasonable assumptions on the analytic properties of the s_{ij}, imply that a knowledge of the reflection coefficient s_{21} (or s_{12}) alone (but not the transmission coefficient alone) determines the remaining coefficients of the S-matrix. The significance of this is discussed in §3.2. Inspection of (1.96) shows that for large k, the effect of the term $V\psi$ is negligible so that

$$s_{11} = s_{22} \sim 1 \quad \text{and} \quad s_{12} \sim s_{21} \sim 0 \quad \text{as} \quad |k| \sim \infty \ . \tag{1.104}$$

1.9.2 The Chandrasekhar Transform

If the HE (1.26) could be transformed into the SE (1.102) the solutions to both direct and inverse scattering problems formulated for the SE could be applied to the HE. Such a transformation exists and has been called the Chandrasekhar Transform (Bates and Wall, 1976). Consider an inhomogeneity which is confined to the half-space $x > 0$ so that

$$\left.\begin{aligned} \nu &= 1 \quad , \quad x \leq 0 \\ &= \nu(x) \quad , \quad x > 0 \ . \end{aligned}\right\} \tag{1.105}$$

The Chandrasekhar transform is motivated by the first order WKB approximate solution (Heading, 1962), denoted here by $\psi \simeq \psi_{WKB}$, to (1.26) which is

$$\psi_{WKB} = \nu^{\frac{1}{2}} \exp(-ik \int_0^x \nu(\xi)\, d\xi) \quad . \tag{1.106}$$

A new coordinate u is defined by

$$\left.\begin{aligned} u &= x \qquad , \quad x \leq 0 \\ &= \int_0^x \nu(\xi)\, d\xi \quad , \quad x > 0 \end{aligned}\right\} \tag{1.107}$$

so that

$$du = \nu dx \tag{1.108}$$

and a new wavefunction ϕ is defined by

$$\phi = \nu^{\frac{1}{2}} \psi \, . \tag{1.109}$$

Equation (1.107) defines the <u>electrical</u> (or optical or acoustic) <u>distance</u> u into the scattering region, and the wavefront travels with constant velocity c with respect to u. Substituting (1.108) and (1.109) into (1.26) gives

$$\partial^2\phi/\partial u^2 + k^2\phi - q\phi = 0 \tag{1.110}$$

which is a SE where the "potential" $q = q(u)$ is given by

$$\begin{aligned} q &= 0 \qquad\qquad\qquad\qquad , \quad u \leq 0 \\ &= (2\nu\, \partial^2\nu/\partial u^2 - (\partial\nu/\partial u)^2)/4\nu^2 \quad , \quad u > 0 \, . \end{aligned} \tag{1.111}$$

Since ψ is identical to ϕ in the free space region (where ψ is measured) then, if the inverse problem for the SE can be solved, q(u) can be determined from the data. Then $\nu(u)$ can be computed (in principle) by solving the non-linear differential equation (1.111) and finally $\nu(x)$ determined using (1.108). The solution to the inverse problem for the SE is described in the following section.

It is worth noting that a similar "inverse transformation" from the SE to the HE has not been found.

1.9.3 The Gelfand-Levitan Equation

The Gelfand-Levitan (GL) technique is a method for solving exactly the inverse scattering problem for the one-dimensional SE. It has received a considerable amount of attention in the more theoretical literature on inverse scattering; probably because it provides an explicit mathematical solution to this particular problem. However the method is not particularly amenable to numerical computation and so there has been very little reported on the processing of real scattering data. The method itself is an outgrowth of the work of two mathematicians, Gelfand and Levitan (1951), on a particular type of differential equation. Their work has been extended and applied to a number of inverse scattering problems. The technique has also been widely used in the solution of non-linear differential equations (Sabatier, 1978).

Most reported derivations of the GL equation use spectral techniques (see, for example, Newton, 1966, §20.2), however these are not easily understood by non-mathematicians unfamiliar with spectral theory. For this reason a more physically motivated derivation in the time domain using causality (following Kay, 1960) is given here. Almost all derivations of the GL equation in the literature gloss over the details (some of which are quite intricate) so I consider it useful to present a fairly detailed treatment here. For convenience, subscripts are used in this section to denote partial derivatives.

Consider the scattering by an inhomogeneous half space $x > 0$ (see Fig. 1.4 (a)) described by the SE (1.96) so that the potential V satisfies

$$V(x) = 0 \quad , \quad x < 0 \; . \tag{1.112}$$

The reflection coefficient from the left, $s_{21}(k)$, which is denoted here by $r(k)$, is measured as a function of k at some point in the free space region $x < 0$. Let $R(x+ct)$ be the reflected field in the region $x < 0$ due to the incident field $\delta(x-ct)$ so that the total field in this region is given by

$$\Psi(x,t) = \delta(x-ct) + R(x+ct) \quad , \quad x < 0 \; . \tag{1.113}$$

Since no reflected field appears before the incident field strikes the scattering region at $x = 0$, causality implies that

$$R(x+ct) = 0 \quad , \quad x < -ct \; . \tag{1.114}$$

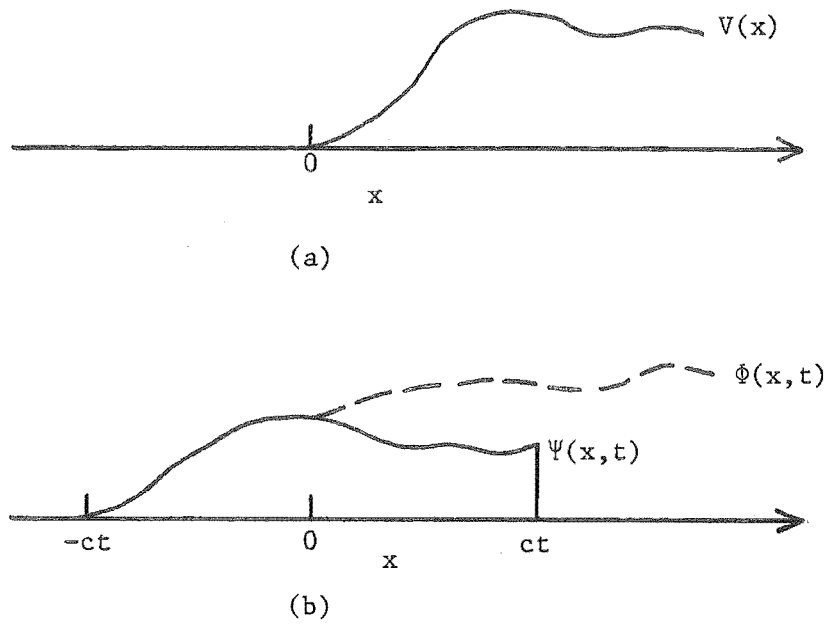


Fig. 1.4 One-dimensional scattering by a potential.
(a) The potential V.
(b) The form of the wavefunctions Φ and Ψ at time t.

The impulse response $R(\tau)$ is equal to the FT of $r(k)$, i.e.

$$R(\tau) = \int_{-\infty}^{\infty} r(k) \exp(-i2\pi k\tau)\, dk \ . \tag{1.115}$$

So $R(\tau)$ may be measured directly or computed from a measurement of $r(k)$. A free space field $\Phi = \Phi(x,t)$ is defined <u>for all x</u> by

$$\Phi(x,t) = \delta(x-ct) + R(x+ct) \quad , \ \forall x \tag{1.116}$$

and hence satisfies the free space equation <u>for all x</u>

$$\Phi_{xx} - (1/c^2)\, \Phi_{tt} = 0 \quad , \quad \forall x \ . \tag{1.117}$$

Hence Φ is an extension of Ψ for $x < 0$ given by (1.113) into the region $x > 0$ assuming that all of space is free ($V = 0$).

An essential step is to assume that the solution Ψ to (1.97) is given by a linear transformation of the solution Φ of the free space equation (1.117) so that

$$\Psi(x,t) = \Phi(x,t) + \int_{-\infty}^{\infty} K(x,z)\ \Phi(z,t)\ dz \quad , \quad x > 0 \ . \tag{1.118}$$

The kernel $K = K(x,z)$ is called the transformation kernel since it transforms a solution of the unperturbed equation (1.117) to a solution of the perturbed equation (1.97). Since the wavefront of Ψ travels with a velocity c then it satisfies the causality relations

$$\Psi(x,t) = 0 \quad , \quad x \leq -ct \quad \text{and} \quad x > ct \ . \tag{1.119}$$

The relations (1.104) indicate that $\Psi(x,t)$ is continuous at $x = -ct$ but discontinuous at $x = ct$, as illustrated in Fig. 1.4(b), because the high frequencies are not reflected, which accounts for the closed and open end-points in (1.119). A detailed description of the propagation of transients in a dispersive medium is given by Karbowiak (1957). It seems reasonable to assume that Ψ will only depend on Φ in the space interval for which Ψ is non-zero, so that

$$K(x,z) = 0 \quad , \quad z \leq -x \quad \text{and} \quad z > x \ . \tag{1.120}$$

Making use of (1.118) and (1.120) gives

$$\Psi(x,t) = \Phi(x,t) + \int_{-x}^{x} K(x,z)\ \Phi(z,t)\ dz \ . \tag{1.121}$$

Equation (1.120) gives one boundary condition,

$$K(x,-x) = 0 \ , \tag{1.122}$$

for $K(x,z)$ and the other conditions are found by substituting (1.121) into (1.97). Differentiating (1.121) twice with respect to x and using (1.122) gives

$$\Psi_{xx}(x,t) = \Phi_{xx}(x,t) + \int_{-x}^{x} K_{xx}(x,z)\ \Phi(z,t)\ dz + 2\ K_x(x,x)\ \Phi(x,t) + K_x(x,-x)\ \Phi(-x,t) + K(x,x)\ \Phi_x(x,t) \ . \tag{1.123}$$

Differentiating (1.121) twice with respect to t and using (1.117) gives

$$\Psi_{tt}(x,t)/c^2 = \Phi_{tt}(x,t)/c^2 + \int_{-x}^{x} K(x,z)\ \Phi_{zz}(z,t)\ dz$$

and integrating twice by parts gives

$$\Psi_{tt}(x,t)/c^2 = \Phi_{tt}(x,t)/c^2 + \int_{-x}^{x} K_{zz}(x,z)\ \Phi(z,t)\ dz - K_x(x,x)\ \Phi(x,t)$$

$$+ K_x(x,-x)\ \Phi(-x,t) + K(x,x)\ \Phi_x(x,t)\ . \qquad (1.124)$$

Substituting (1.121), (1.123) and (1.124) into (1.97) and using (1.117) gives

$$\int_{-x}^{x} (K_{xx} - K_{zz} - VK)\ \Phi\ dz + (2K_x(x,x) - V(x))\ \Phi = 0 \quad , \quad x > 0$$

which requires that $K(x,z)$ satisfies the differential equation

$$K_{xx} - K_{zz} - VK = 0 \qquad (1.125)$$

and the boundary condition

$$V(x) = 2K_x(x,x)\ . \qquad (1.126)$$

Finally, substituting (1.116) into (1.121) and using (1.114) and (1.119) gives

$$R(x+ct) + K(x,ct) + \int_{-ct}^{x} K(x,z)\ R(z+ct)\ dz = 0 \quad , \quad x > ct \qquad (1.127)$$

which is called the Gelfand-Levitan equation. The GL equation must be solved as a preliminary to solving the inverse scattering problem, i.e. $R(\tau)$ is computed from $r(k)$ and the integral equation (1.127) is solved for $K(x,z)$, for each value of x (regarded as a parameter), considered as a function of z over the range $-x < z < x$. The potential V is then given by (1.126).

If the potential has bound state solutions (these correspond to poles of $r(k)$ on the positive imaginary k-axis) then additional terms must be included on the RHS of (1.115) to take account of these (see Chadan and Sabatier, 1977, ch. 2). Kay (1960) shows that if $r(k)$ is a rational function (characterised by a finite number of poles and zeros), i.e. of the form

$$r(k) = A \prod_{n=1}^{N} (k-\mu_n) \Big/ \prod_{m=1}^{M} (k-\lambda_m) \qquad (1.128)$$

where A and the μ_n and λ_m are constants, then a closed form expression for $V(x)$ can be obtained.

Note that in a quantum mechanical inverse scattering problem (just as in X-ray crystallography) only $|r(k)|^2$ can be measured experimentally

and the phase problem must be solved (Newton, 1966, §20.1) before the GL technique can be applied. The technique of extending the free space field into the scattering region and then requiring that the total field satisfy causality conditions is similar, in a sense, to the "null field" method (Bates and Wall, 1977) used to solve the direct scattering problem for scatterers with sharp boundaries.

Extensions of the GL technique to three space dimensions (Kay and Moses, 1961, Newton, 1974 and Chadan and Sabatier, 1977, ch. 14) have been made, but the practical implications of these are not well understood. Also, the three-dimensional GL method has not yet been adapted to handle the three-dimensional HE, the difficulty arising from trying to generalise (1.107) to three dimensions. Bates and Millane (1981) propose a generalisation of (1.107) to three dimensions, but the significance of this needs further investigation.

1.10 THE INVERSE EIGENVALUE PROBLEM

An inverse scattering problem of interest (particularly in geophysics) is the determination of the properties of a body from its natural frequencies of oscillation. After a large earthquake the earth oscillates for many days, and the frequencies of a number of these modes of oscillation have been measured quite accurately (Gilbert and Dziewonski, 1975). The question arises as to what properties of the earth (for example, the density and elastic parameters as a function of radius) can be determined from this data set. Problems of this type are referred to as inverse eigenvalue or inverse normal mode problems. In this section some of the characteristics of a simple one-dimensional inverse eigenvalue problem are considered.

Equation (1.22) can be cast as an eigenvalue problem by placing appropriate boundary conditions (at, say, $x = 0$ and $x = a$) on ψ. The boundary conditions usually considered are homogeneous so that

$$\psi + \alpha \, \partial\psi / \partial x = 0 \tag{1.129}$$

at the particular boundary and α is a constant. The eigenvalue problem is satisfied by discrete values of k^2, denoted by k_n^2, which are called the eigenvalues. Note that the eigenvalues are proportional to the squares of the resonant frequencies. The corresponding wavefunctions ψ_n, called the modes or eigenfunctions, satisfy

$$\partial^2\psi_n/\partial x^2 + k_n^2\,\nu^2\,\psi_n = 0 \quad . \tag{1.130}$$

The eigenfunctions form a complete orthogonal set with weight function ν^2 on the interval (0,a) - see, for example, Morse and Feshbach (1953) §6.3.

It is convenient to consider two eigenvalue problems on the interval (0,a), with eigenfunctions ψ_n and ϕ_n, defined by the following boundary conditions. Both ψ_n and ϕ_n satisfy the Dirichlet (or sound soft) boundary condition at x = a so that

$$\psi_n(a,k) = \phi_n(a,k) = 0 \quad . \tag{1.131}$$

At x = 0, ψ_n and ϕ_n satisfy Dirichlet and Neumann (or sound hard) boundary conditions respectively so that

$$\psi_n(0,k) = 0 \tag{1.132}$$

and

$$\partial\phi_n(0,k)/\partial x = 0 \quad . \tag{1.133}$$

The eigenvalues of ψ_n and ϕ_n are denoted by k_n^2 and ℓ_n^2 respectively. If $\nu(x)$ is differentiable then it can be shown (Krein, 1952) that the eigenvalues satisfy the asymptotic relations

$$k_n \sim n\pi/A \quad \text{as} \quad n \sim \infty \tag{1.134}$$

and

$$\ell_n \sim (n-\tfrac{1}{2})\pi/A \quad \text{as} \quad n \sim \infty \tag{1.135}$$

for large n where the constant A is the integrated refractive index on the interval (0,a) given by

$$A = \int_0^a \nu(\xi)\,d\xi \quad . \tag{1.136}$$

Hence the average refractive index $\langle\nu\rangle$ on (0,a) is given by

$$\langle\nu\rangle = A/a \quad . \tag{1.137}$$

It can also be shown (Barcilon, 1979) that the k_n and ℓ_n interlace, i.e.

$$0 < \ell_1 < k_1 < \ell_2 < k_2 < \ldots \quad . \tag{1.138}$$

The problem of determining $\nu(x)$ from sets of eigenvalues (these are sometimes called discrete spectra or simply spectra) is a simple example of the inverse eigenvalue problem which has been studied, for instance

by Borg (1946), Krein (1952) and Barcilon (1979). Since (1.22) also describes small transverse displacements on a string of constant tension and mass density $\nu^2(x)$ stretched between $x = 0$ and $x = a$, it is occasionally referred to as the "string problem". A very similar problem (sometimes called the "inverse Sturm-Liouville problem"), the determination of the potential $V(x)$ from eigenvalues of the Schrodinger equation (1.96), has also been studied by Borg (1946), Levitan (1964) and Barcilon (1975). It has been shown (Borg, 1946; Krein, 1952) that $\nu(x)$ is determined uniquely if both the spectra $\{k_n\}$ and $\{\ell_n\}$ are known for all n. The equivalent result applies to the inverse Sturm-Liouville problem (Levitan, 1964). In fact the same result applies if the boundary conditions (1.131) to (1.133) are replaced by any homogeneous boundary conditions that are identical at one endpoint and distinct at the other (Levitan, 1964).

That $\nu(x)$ is uniquely determined by the two spectra can be shown by treating it as a scattering problem with an incident wave from the left. The reflection coefficient is a meromorphic function (regular except for poles) given by (Gerver, 1970)

$$r(k) = A \prod_{n=1}^{\infty} \frac{(1 - k^2/k_n^2)}{(1 - k^2/\ell_n^2)} . \tag{1.139}$$

As a result of the boundary condition (1.131) all of the incident energy is reflected so that

$$|r(k)| = 1 . \tag{1.140}$$

Equation (1.139) shows that the two spectra $\{k_n\}$ and $\{\ell_n\}$ determine the reflection coefficient, and it was shown in §1.9 that the potential or the refractive index is uniquely determined by the reflection coefficient. Hence $\nu(x)$ is uniquely determined by $\{k_n\}$ and $\{\ell_n\}$ for all n. If it is known *a priori* that $\nu(x)$ is symmetric on (0,a), i.e.

$$\nu(x) = \nu(a - x) , \tag{1.141}$$

then only the single spectrum $\{k_n\}$ is required, since ψ and $\partial\psi/\partial x$ are equal to zero alternately at $x = a/2$ and so $\{k_n\}$ is equivalent to two spectra for ν on the interval (0,a/2).

Usually the boundary conditions are determined by the physics of the particular problem being studied. Hence, often, only one spectrum is available for measurement and this provides insufficient data for a unique

solution. However some other information may be equivalent to the unavailable spectrum. For example, Levitan (1964) has shown that one spectrum plus the slope of the eigenfunctions at one endpoint provide a sufficient data set for a unique solution of the inverse Sturm-Liouville problem. The same problem occurs for the normal modes of oscillation of the earth because they can only be measured for one boundary condition at the surface. Barcilon (1978) has examined the information content of the normal modes of a spherically symmetric fluid sphere associated with different angular orders but it appears that such a data set does not uniquely determine the velocity profile of such a sphere.

1.11 THE BACKUS AND GILBERT METHOD

In this section an important model fitting type procedure of very general application is described. The method is due to Backus and Gilbert (1968) who have applied it to a number of geophysical inverse problems; for example the inverse normal mode problem described in §1.10. It has also found wide application in atmospheric profile inversion (Rodgers, 1976). The method can be applied to the determination of a number of profile functions but for simplicity, the case of a single profile is considered here.

Firstly, the linear inverse problem is considered where the profile $f(x)$ is to be estimated from a set of N linearly independent measurements m_i, which are linear functionals of the profile so that

$$m_i = \int_L g_i(x)\, f(x)\, dx \quad , \quad i = 1,2,\ldots,N \tag{1.142}$$

where the $g_i(x)$ are N known functions and L is the interval over which $f(x)$ exists. Normally $f(x)$ will be subject to physical restrictions or *a priori* constraints (for example a mass density must be positive) which are usually in the form of inequalities. Each m_i could be a single normal mode, for instance, or perhaps a satellite radiance measurement through the atmosphere in a single spectral band. The inverse problem is to determine the profile $f(x)$ from the data set $\{m_i\}$. Equation (1.142) is a Fredholm integral equation of the first kind (Porgorzelski, 1966) for $f(x)$, the solution to which is unstable or "improperly posed" (as described in §1.1) in that small errors in the data m_i can lead to large errors in $f(x)$. A common method of stable solution of equations of this

type is regularisation (cf. Deschamps and Cabayan, 1972). However, if N is so small that the solution cannot be stabilised, then the original problem is not directly solvable. In this case all that can be reasonably done is to take a linear combination of the measurements m_i. This gives a linear function of f(x). The essence of the Backus and Gilbert method is to control the shape of this function to provide meaningful information on f(x).

Let the estimate of f(x) at x = y be h(y) which is determined from a linear combination of the data so that

$$h(y) = \sum_{i=1}^{N} a_i(y)\, m_i \ . \tag{1.143}$$

Substituting (1.142) into (1.143) gives

$$h(y) = \int_L A(x,y)\, f(x)\, dx \tag{1.144}$$

where

$$A(x,y) = \sum_{i=1}^{N} a_i(y)\, g_i(x) \ . \tag{1.145}$$

It is convenient to normalise A(x,y) so that

$$\int_L A(x,y)\, dx = 1 \ . \tag{1.146}$$

The ideal form of A(x,y) is $\delta(x-y)$ since then h(y) = f(y) would provide a perfect solution. The idea is to choose the $a_i(y)$ so that A(x,y) resembles $\delta(x-y)$ as closely as possible, under the constraint (1.146), so that (1.144) gives a local average of f(x) around the point x = y with as good resolution as possible. This is done by defining the "spread" or "non-deltaness", S(y), of A(x,y) around the point x = y by

$$S(y) = \int_L J(x,y)\, A(x,y)\, dx \tag{1.147}$$

where J(x,y) is small near x = y and large away from x = y. Backus and Gilbert have found $J(x,y) = (x-y)^2$ to be reasonably adequate. The spread, S(y), is minimised with respect to the $a_i(y)$ under the constraint (1.146) to find the best values for $a_i(y)$. This is a constrained optimisation problem and can be solved using Lagrange multipliers (Backus and Gilbert, 1968). This process must be repeated for each value of y where the profile is to be estimated. Once the $a_i(y)$ are determined, the estimate h(y) of the profile is computed using (1.143). The resolution in the reconstructed

profile can be estimated by computing A(x,y) as illustrated in Fig. 1.5(a).

In practice, of course, account must be taken of the noise in the measurements and so a weighted sum of the spread and the noise variance is minimised. The relative weights assigned to the spread and the noise are chosen according to the application to give the best tradeoff between resolution and noise sensitivity as illustrated in Fig. 1.5(b).

If the measurements are non-linear functionals of f(x) then an initial estimate $\tilde{f}(x)$ of f(x) is chosen and the data $\tilde{m}_i$ it would produce are computed. Hopefully $\tilde{f}(x)$ is sufficiently close to f(x) that the problem can be linearised so that, to a reasonable approximation, $(m_i - \tilde{m}_i)$ is a linear functional of $(f(x) - \tilde{f}(x))$. Using the above technique, an improved estimate of f(x) can be computed to replace $\tilde{f}(x)$. The process is repeated in an iterative fashion until there is no significant change in $\tilde{f}(x)$.

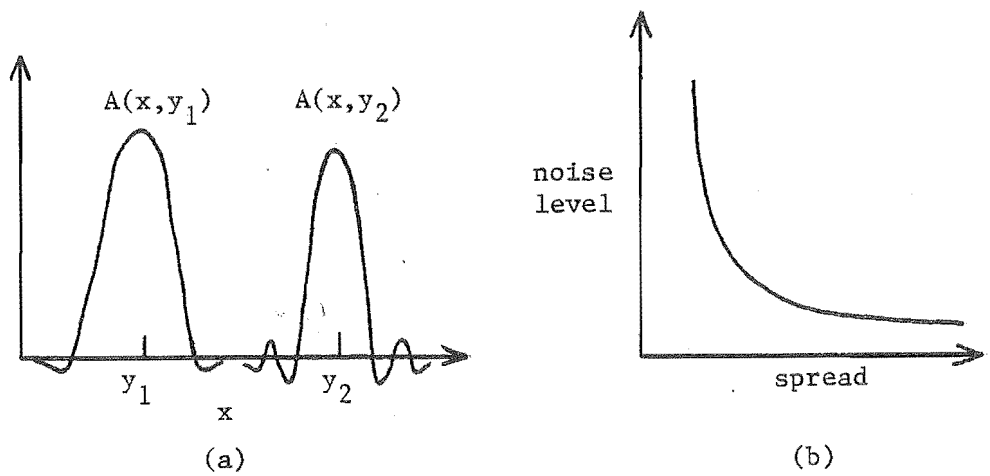


Fig. 1.5 The Backus and Gilbert Method.
(a) Form of the weighting functions A(x,y) for different values of y.
(b) Form of the resolution-noise tradeoff curve.

1.12 SUMMARY

The topics discussed in this chapter serve as a summary of the main inverse scattering techniques for penetrable bodies and as background to chapters 2 and 3. The methods of X-ray crystallography outlined in §1.7 are extended and applied to an analysis of DNA structure in chapter 2.

Chapter 3 is primarily concerned with exact solutions to inverse scattering for the HE in stratified media and in branched ducts and transmission lines. Since chapter 3 is concerned with macroscopic wave motion described by the HE, the GL method does not directly apply so that alternative time domain procedures are adopted and their relationship to the GL method is discussed. These techniques are compared to the Born approximation outlined in §1.5 and to a first order solution to the inverse eigenvalue problem introduced in §1.10.

2. THE STRUCTURE OF DNA

2.1 INTRODUCTION

Deoxyribonucleic Acid (DNA) is one of the most important molecules in the living cell (Watson, 1976). It controls cell processes by regulating the production of proteins, and "stores" genetic information. DNA is found within the chromosomes (made up of DNA and protein) of cells where all the genetic information of the organism, to which the cell belongs, is stored. It is a large molecule (since it stores an enormous amount of genetic information) and belongs to the class of so-called biological macromolecules. The DNA molecule also enables the genetic information to be transferred during cell reproduction. The total length of DNA in a human cell is about one metre and contains about 10^{11} atoms. Thus long lengths of DNA are folded up so that they fit into the cell. The detailed molecular structure of DNA has far reaching implications in (a) the study of how it packs into the cell and (b) the mechanisms for protein synthesis and genetic code transcription.

DNA is a long chain-like molecule, and the individual units which link together to form the chain are called nucleotides. Molecules of this type are called polynucleotides. A nucleotide consists of three groups of atoms called a phosphate, a sugar (deoxyribose) and a base, as shown in Fig. 2.1(a). The base is one of four called adenine, thymine, guanine and cytosine, abbreviated to A, T, G and C respectively. The nucleotides link together to form a strand as shown in Fig. 2.1(a). The part of the strand made up of the sugars and phosphate groups is called the sugar-phosphate backbone. It is primarily the geometrical association of these strands in DNA which is investigated in this chapter.

Watson and Crick (1953) proposed the now familiar double helix (DH) model for DNA. By building models of the nucleotides and deriving clues from the X-ray diffraction pattern, they produced the DH model shown in Fig. 2.1(b). This structure is made up of two strands running in a right-handed helical fashion around the molecular axis. The two strands are joined together by their bases which form base pairs as shown in Fig. 2.1. The two bases of a base pair are linked together by hydrogen bonds (relatively weak forms of chemical bonds). Two nucleotides

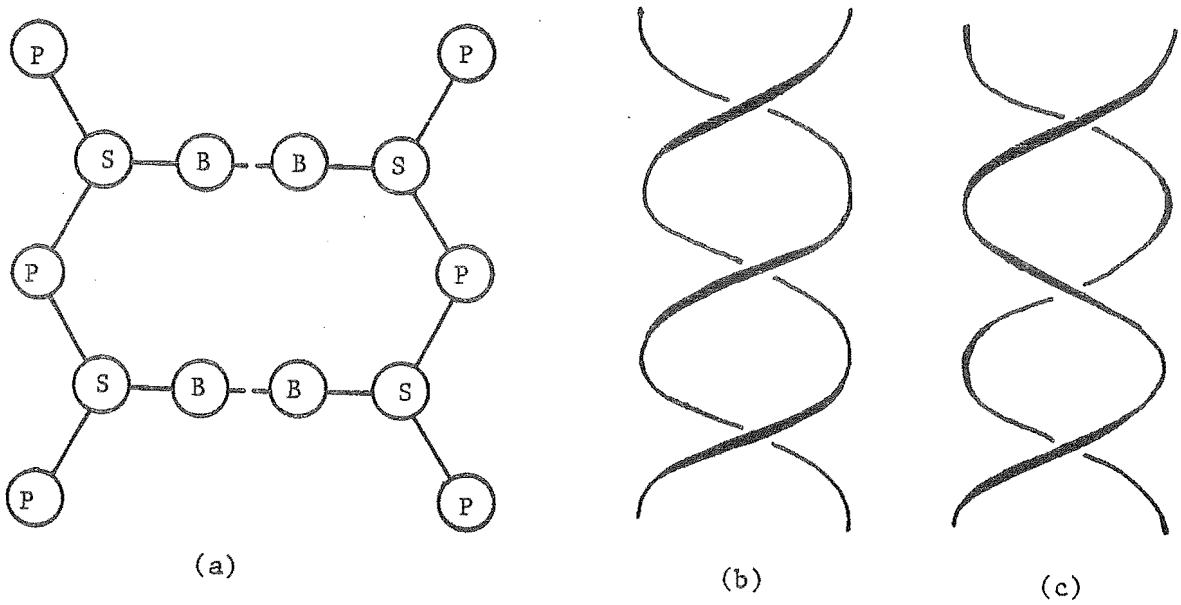


Fig. 2.1 The geometrical arrangement of DNA.
(a) The arrangement of the phosphate (P), sugar (S) and base (B) groups.
(b) The sugar-phosphate strands of the DH model.
(c) The sugar-phosphate strands of the SBS model.

containing a single base pair are referred to here as a nucleotide pair. The two strands are antiparallel, which means that they have equivalent structure but run in opposite directions. Probably the most important feature of Watson and Crick's structure is that the bases pair together only in specific combinations - A with T and G with C. It is the sequence of the bases along the chain which codes the genetic information. The specific base pairing explains how the genetic information is transferred when a cell divides. The two strands separate (denature) when the cell divides and one strand goes into each daughter cell. The specific base pairing allows the correct replacement strands to be built onto each of these single strands forming two new DNA molecules, each containing the original genetic information. However the DH has a major shortcoming with respect to this denaturing process. Because of the extensive intertwining of the two strands, they must rotate relative to each other in order to separate. Taking into account the time taken for denaturation and the length of the molecule, it is found that they must rotate extremely rapidly (Bates *et al.*, 1977). The mechanism by which this separation is thought to occur so rapidly without the strands getting tangled is extremely intricate.

A number of enzymes are known to be involved *in vivo* (in the living state) but the manner in which they assist "unwinding" is not known. In addition to the unwinding problem, the agreement between the observed X-ray diffraction data and that predicted by the DH is not all that good, even after 20 years of refinement. It is important to realise that it is the base pairing scheme proposed by Watson and Crick (and not the double helical disposition of the two strands) which has had such a profound influence on biology.

In 1976, Gordon Rodley of the Chemistry department of this University proposed the side-by-side (SBS) model as a possible alternative structure for DNA (Rodley *et al.*, 1976). This model retains the Watson-Crick base pairing scheme with two antiparallel sugar-phosphate strands on the outside of the molecule. However the two strands twist one way and then the other and so lie, essentially, side by side (see Fig. 2.1(c)). The SBS model can be considered, approximately, as a succession of alternating left and right-handed double helical segments. The major appeal of the SBS structure is that the two strands can more simply pare apart in replication, and other situations. In order to assess the viability of the SBS structure, it is essential to determine if it is compatible with the observed X-ray diffraction data. A number of people in this department - R.H.T. Bates, C.H. Rowe, R.M. Lewitt, P.M. Andreae, G.C. McKinnon and myself - have contributed in this regard. In this chapter a number of increasingly accurate representations of the SBS and DH models are compared with the observed X-ray diffraction data. The stereochemical characteristics of the SBS model are of prime importance in determining the structure's viability. Stereochemical refinement of the structure is also described in this chapter.

The two chief forms in which DNA crystallises (depending on the relative humidity of the sample) are called A-DNA and B-DNA. The analysis presented in this chapter applies to B-DNA. B-DNA forms at 92% relative humidity and the unit cell is orthorhombic. X-ray diffraction data shows that B-DNA repeats along its axis (forming a repeat unit of length 33.8 Å) with 10 nucleotide pairs per repeat unit, and the base pairs are approximately perpendicular to the molecular axis. The high relative humidity of DNA crystals means that there is a large quantity of water dispersed between the molecules. Because the sequence of base pairs is variable, the molecule is not strictly periodic along its axis. However the base pairs are similar in structure and so the molecule can be

considered periodic with "average bases". The two strands are labelled A and B and the nucleotide pairs of a repeat unit are numbered 1 to 10. The z-axis is chosen to coincide with the molecular axis and the magnitudes a, b and c of the lattice vectors are equal to 30.76 Å, 22.50 Å and 33.80 Å respectively.

Crystalline samples of DNA so far obtained are so poor that the quality of the measured X-ray diffraction pattern is not sufficiently high to allow unambiguous structure determination to be made using the techniques outlined in §1.7. The direct interpretation of diffraction patterns in terms of molecular structure is difficult because of the FT relationship described in §1.7. The effect of an atom is spread throughout all of reciprocal space and the intensity at a point in reciprocal space is due to the effects of all the atoms in the molecule. As described in §1.7, the Patterson function is calculable directly from the diffraction data and is (in principle) more easily interpreted in terms of molecular structure than the diffraction pattern. For this reason the analysis presented in this chapter is in terms of Patterson-type functions. The low quality of the X-ray data means that a conventional Patterson cannot be calculated and so alternative Patterson-type functions are used to compare model structures with the data. Because the phosphate groups are strong scatterers of X-rays, initial comparisons with the X-ray data concentrated on the effects of these groups. Bates *et al.* (1977) showed that a SBS model consisting of filamentary strands of electron density following the sugar-phosphate backbone gives a diffraction pattern in general agreement with that observed. They also showed that a filamentary model of the DH is not consistent with certain features of the observed diffraction pattern.

In §2.2 and §2.3 the SBS and DH models are compared (using the phosphate groups only) in terms of the paracrystalline and fibre diffraction data respectively. This work was followed up by Graeme McKinnon (1980) who computed "axial Pattersons" and diffraction patterns using all the atoms (except the hydrogen atoms since they are such weak scatterers that their effect is insignificant) in the molecule. The results of this work are reported in Bates *et al.* (1978) and Bates *et al.* (1980). An accurate physical model of the SBS structure was built by Gordon Rodley, and the atomic coordinates measured. In §2.4 the stereochemical refinement of these coordinates to bring the model within acceptable stereochemical limits is described. The stereochemically refined coordinates and other stereochemical details are reported in Millane and Rodley (1981) and Millane *et*

al. (1982). Axial Pattersons of this stereochemically refined structure are described in §2.5 and the significance of this work is discussed in §2.6.

2.2 PARACRYSTALLINE ANALYSIS

Many macromolecular specimens cannot be prepared as good single crystals. Paracrystalline specimens of DNA consist of molecules which have all their molecular axes parallel but are made up of crystallites randomly rotated with respect to each other. Each crystallite consists of molecules forming a perfect crystal as described in §1.7.2. Because of the presence of crystallites, the diffraction pattern is observed only at the reciprocal lattice points. However the random orientation of the crystallites means that all structure intensities $|F_{hk\ell}|^2$ with the same value of $(h^2 + k^2)^{\frac{1}{2}}$ are superimposed. Hence, taking (1.75) into account, the independent measurable quantities are $|F|^2_{hk\ell}$ defined by

$$|F|^2_{hk\ell} = (\varepsilon_h \varepsilon_k \varepsilon_\ell / 8)\ (|F_{hk\ell}|^2 + |F_{\bar{h}k\ell}|^2 + |F_{h\bar{k}\ell}|^2 + |F_{\bar{h}\bar{k}\ell}|^2 + |F_{hk\bar{\ell}}|^2 + |F_{\bar{h}k\bar{\ell}}|^2 + |F_{h\bar{k}\bar{\ell}}|^2 + |F_{\bar{h}\bar{k}\bar{\ell}}|^2) \qquad (2.1)$$

where ε_n is the Neumann factor defined by

$$\left.\begin{aligned} \varepsilon_n &= 1 \quad , \quad n = 0 \\ &= 2 \quad , \quad n \neq 0 \end{aligned}\right\} \qquad (2.2)$$

and h, k and ℓ are non-negative integers.

The unit cell contains one molecule of DNA in the centre and one at each of the four corners as shown in Fig. 2.2. Let $g = g(x,y,z)$ and $f = f(x,y,z)$ be the electron densities of the crystal and a crystal which contains only the molecules in the centres of the unit cells respectively so that

$$g(x,y,z) = f(x,y,z) + f(x - a/2,\ y - b/2,\ z - cd) \qquad (2.3)$$

where d is the fractional shift along the z-axis of the corner molecules relative to the central molecules. Taking the FT of (2.3) gives

$$G(u,v,w) = F(u,v,w)\ [1 + \exp(i2\pi(ua/2 + vb/2 + wcd))]$$

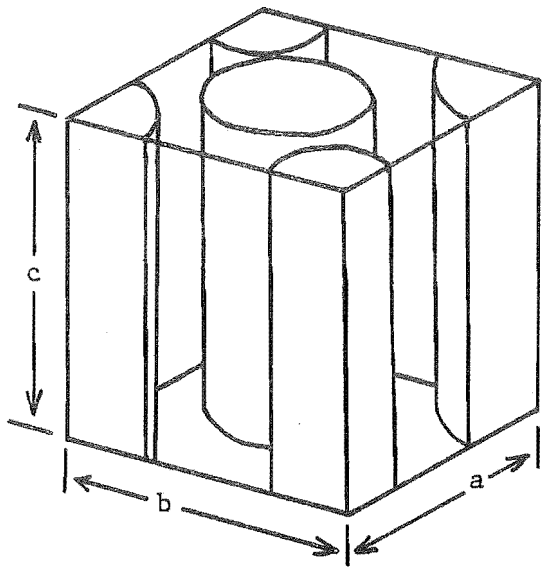


Fig. 2.2 The contents of the unit cell.

so that

$$|G_{hk\ell}|^2 = 2\ |F_{hk\ell}|^2\ [1 + \cos(\pi(h+k+2\ell d))] \qquad (2.4)$$

where upper case letters denote the FT. Making use of (2.1) and (2.4) gives

$$|F|^2_{hk\ell} = |G|^2_{hk\ell}/[4\ \cos^2(\pi(h+k+2\ell d)/2)] \qquad (2.5)$$

so that the measured data $|G|^2_{hk\ell}$ can be transformed to that which would be obtained if there was only one molecule per unit cell. Making use of this transformation reduces the number of intermolecular peaks in Pattersons calculated from the data, simplifying their interpretation. Note that this transformation can only be made when the denominator in (2.5) is non-zero so that

$$h + k + 2\ell d \neq 2n + 1 \qquad (2.6)$$

where n is an integer. For A-DNA d = 0 so that only half the values of $|F|^2_{hk\ell}$ can be obtained from the data. For B-DNA d = 1/3 so that $|F|^2_{hk\ell}$ can be obtained for all values of h and k if ℓ is neither zero nor a multiple of three.

Since the individual structure intensities cannot be measured, the Patterson, P (as defined in §1.7.3), cannot be computed from the data. What can be computed is the cylindrically averaged Patterson $\tilde{P} = \tilde{P}(r,z)$ (MacGillavry and Bruins, 1948) defined by

$$\tilde{P}(r,z) = (1/2\pi) \int_0^{2\pi} P(r\cos\theta,\ r\sin\theta,\ z)\ d\theta \qquad (2.7)$$

where (r,θ,z) are cylindrical polar coordinates. This is related to the data by

$$\tilde{P}(r,z) = (1/abc) \sum_{h,k,\ell=0}^{\infty} |F|^2_{hk\ell}\ J_o(2\pi r(h^2/a^2 + k^2/b^2)^{\frac{1}{2}})\ \cos(2\pi\ell z/c) \qquad (2.8)$$

It is convenient to leave the evaluation of (2.8) until the discussion of fibre patterns in §2.3.

In this section, different models of the structure of DNA are assessed by examining the dependence of Patterson type functions on coordinates transverse to the molecular axis (i.e. in the x-y or r-θ plane). Because $\tilde{P}(r,z)$ is equal to the Patterson averaged over θ, its r dependence is not easy to interpret. A new Patterson type function, whose dependence on x and y is easier to interpret in terms of molecular structure, is introduced in this section. Note however that the axial or z dependence of $\tilde{P}(r,z)$ is very informative, and extensive use is made of this in §2.3 and §2.5.

Making use of (1.73) allows the structure intensities to be written as

$$|F_{hk\ell}|^2 = \iiint_{-\infty}^{\infty} A(x,y,z)\ \exp(i2\pi(hx/a + ky/b + \ell z/c))\ dxdydz \qquad (2.9)$$

where $A(x,y,z)$ is the autocorrelation of a single unit cell defined by (1.79). Using (2.1) and (2.9) gives

$$|F|^2_{hk\ell} = \iiint_{-\infty}^{\infty} A(x,y,z)\ \cos(2\pi hx/a)\ \cos(2\pi ky/b)\ \cos(2\pi\ell z/c)\ dxdydz\ . \qquad (2.10)$$

The function $\Phi_\ell(x,y)$ (cf. Bates *et al.*, 1978) is defined by

$$\Phi_\ell(x,y) = (1/4ab) \sum_{h,k=o}^{\infty} |F|^2_{hk\ell}\ \cos(2\pi hx/a).\ \cos(2\pi ky/b)\ . \qquad (2.11)$$

Substituting for $|F|^2_{hk\ell}$ from (2.10) in (2.11) gives

$$\Phi_\ell(x,y) = (1/4) \int_{-\infty}^{\infty} [\hat{P}(x,y,z) + \hat{P}(-x,y,z) + \hat{P}(x,-y,z) + \hat{P}(-x,-y,z)]\cos(2\pi\ell z/c)dz \qquad (2.12)$$

where the function $\hat{P}(x,y,z)$, which takes the form of a Patterson in the x and y directions and an autocorrelation in the z direction, is given by

$$\hat{P}(x,y,z) = \sum_{h,k=-\infty}^{\infty} A(x - ha,\ y - kb,\ z)\ . \qquad (2.13)$$

Since $\hat{P}$ is symmetric about the origin, (2.12) can be used to relate $\Phi_\ell(x,y)$ to Pattersons in the x and y directions so that

$$\Phi_\ell(x,0) = \int_{-\infty}^{\infty} \hat{P}(x,0,z) \cos(2\pi\ell z/c)\, dz \qquad (2.14)$$

and

$$\Phi_\ell(0,y) = \int_{-\infty}^{\infty} \hat{P}(0,y,z) \cos(2\pi\ell z/c)\, dz \ . \qquad (2.15)$$

Hence the functions $\Phi_\ell(x,0)$ and $\Phi_\ell(0,y)$, henceforth called the Phi functions, are Pattersons in the x and y directions respectively, together with a weighted average in the z direction. The Phi functions can be computed from the observed data using (2.11) and from model structures using (2.14) and (2.15). Inspection of (2.13) to (2.15) shows that the Phi functions do not involve averaging of the Patterson in the x-y plane and so are suitable for comparing the SBS and DH structures.

The Phi functions are computed using the diffraction data for paracrystalline B-DNA listed in Arnott and Hukins (1973). The measured data $|G|^2_{hk\ell}$ are reduced to the $|F|^2_{hk\ell}$, which corresponds to one molecule per unit cell, using (2.5) for $\ell = 1$ and $\ell = 2$. In practice the $|F|^2_{hk\ell}$ are recorded as diffraction spots on film. Some of these diffraction spots overlap so that not all of the individual $|F|^2_{hk\ell}$ can be obtained. Table 2.1 shows the number of the $|F|^2_{hk\ell}$ which can be obtained uniquely and those which can be obtained if, in addition, the ones which overlap in pairs in Table 1 of Arnott and Hukins (1973) are also included.

Table 2.1 The number of values of $|F|^2_{hk\ell}$ used to compute $\Phi_\ell(x,0)$ and $\Phi_\ell(0,y)$ in Figs 2.3 and 2.4.

ℓ	Single spots	Single and double spots
1	16	28
2	9	16

Fig. 2.3 shows the Phi functions computed using the $|F|^2_{hk\ell}$ obtained from the single spots only. For either x or y less than about 2 Å, the Phi functions $\Phi_\ell(x,0)$ and $\Phi_\ell(0,y)$ are proportional to the autocorrelations of the structure in the x and y directions respectively. Since the atoms, to

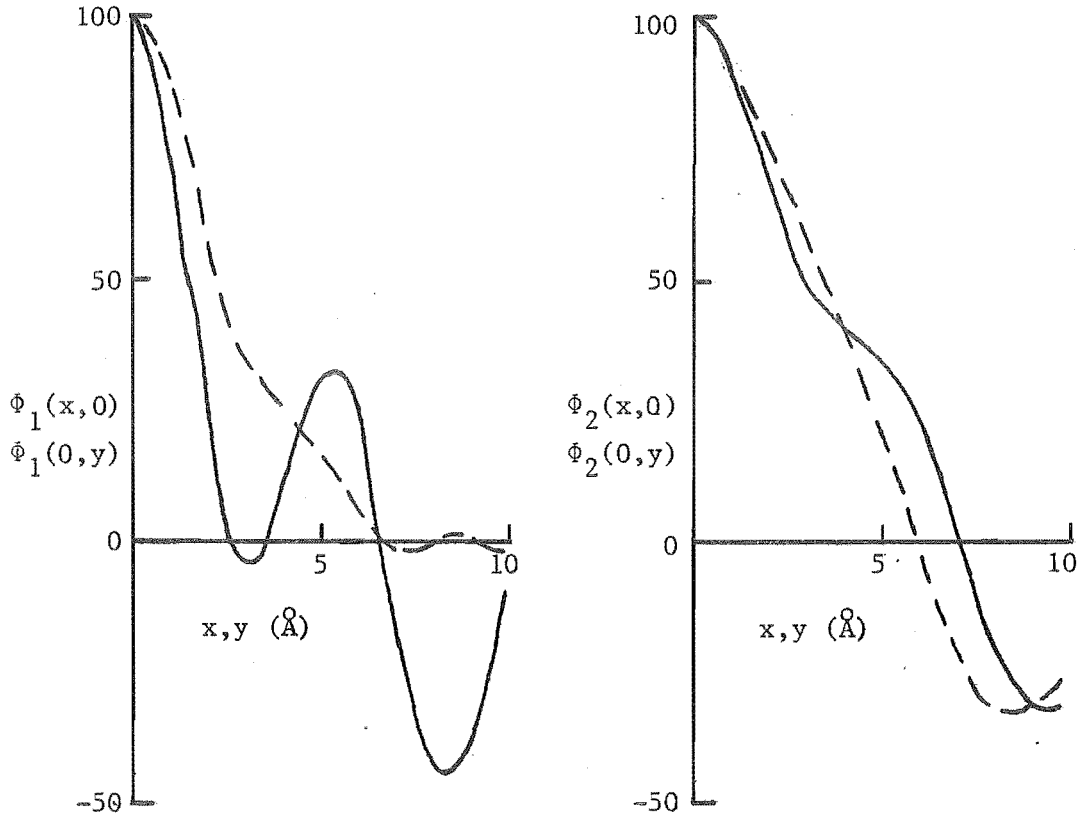


Fig. 2.3 Phi functions $\Phi_\ell(x,0)$ (---) and $\Phi_\ell(0,y)$ (——) calculated from the diffraction data (single spots only) given in table 1 of Arnott and Hukins (1973).

a good approximation, are spherically symmetric the autocorrelation is independent of direction so that

$$\Phi_\ell(\xi,0) \simeq \Phi_\ell(0,\xi) \quad , \quad \xi \leq 2\,\text{Å} \; . \tag{2.16}$$

Inspection of the curves in Fig. 2.3 shows that (2.16) is not satisifed, suggesting that the resolution has been limited by the small number of the $|F|^2_{hk\ell}$ used. Fig. 2.4 shows $\Phi_\ell(x,0)$ and $\Phi_\ell(0,y)$ computed using the $|F|^2_{hk\ell}$ obtained from the single and double spots. The intensity recorded for the double spots was divided equally (an arbitrary but reasonable choice) between the two $|F|^2_{hk\ell}$. The curves in Fig. 2.4 show a clear improvement in resolution since (2.16) is now satisfied quite well. The main features of the curves in Figs 2.3 and 2.4 are the same, so (despite the arbitrary apportioning of the $|F|^2_{hk\ell}$ for the double spots) those in Fig. 2.4 are probably reasonably accurate. The curves for $\ell = 2$ have very little structure indicating that the resolution is quite low. Hence only the curve for $\ell = 1$ in Fig. 2.4 is used to compare the data with different models of DNA.

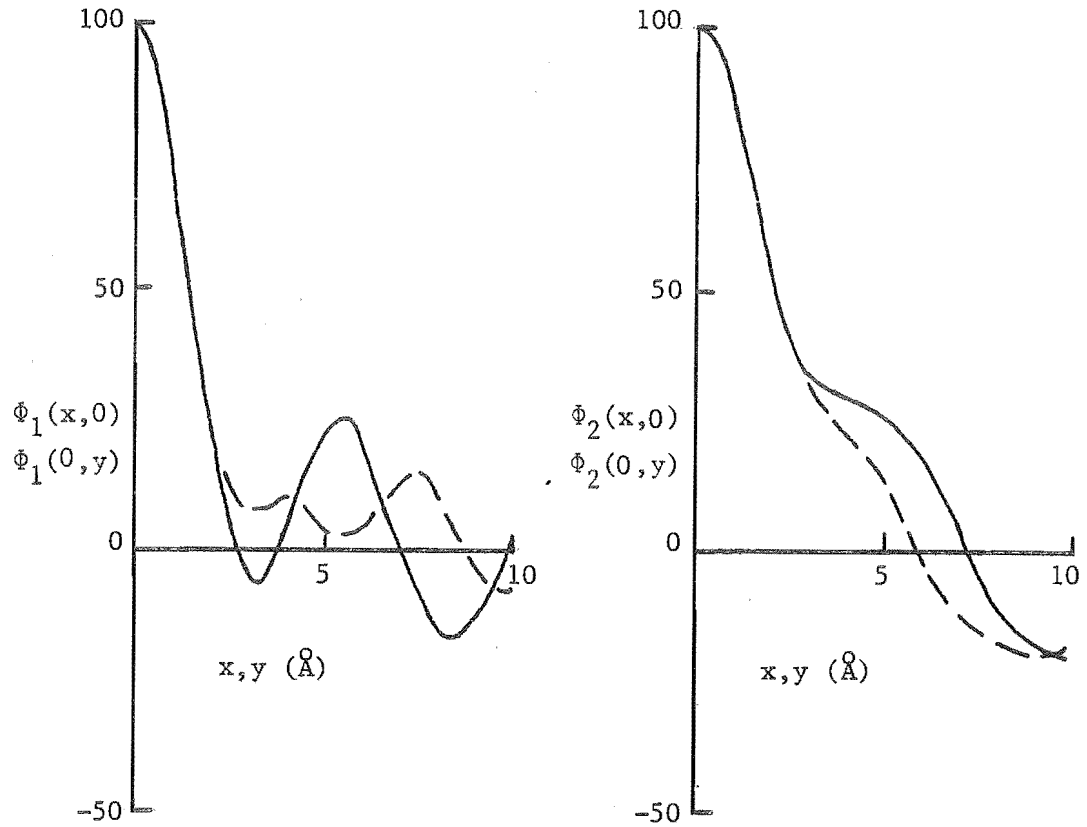


Fig. 2.4 Phi functions $\Phi_\ell(x,0)$ (---) and $\Phi_\ell(0,y)$ (——) calculated from the diffraction data (single and double spots) given in table 1 of Arnott and Hukins (1973).

In order to compare the SBS and DH models of DNA with the diffraction data, $\Phi_1(x,0)$ and $\Phi_1(0,y)$ are computed for the two models. To extend the work of Bates *et al.* (1977) (who represented the different DNA models as continuous strands of electron density following the sugar-phosphate backbones) to more accurate representations to which modifications can be made, it was considered appropriate to use a model consisting of the phosphate groups only. This allows the atomic coordinates to be adjusted to improve the agreement with the X-ray data. Since a phosphate group is symmetric about its phosphorus atom, this is the same as considering the phosphorus atoms only. Because of the different shapes of the backbones in the SBS and DH structures, the positions of the phosphorus atoms are quite different (as shown in projection onto the x-y plane in Fig. 2.5) allowing a comparison of the two models. Each phosphorus atom is approximated by a Gaussian ball of electron density $\rho(r)$ so that

$$\rho(r) = \exp(-r^2/2\ r_o^2) \qquad (2.17)$$

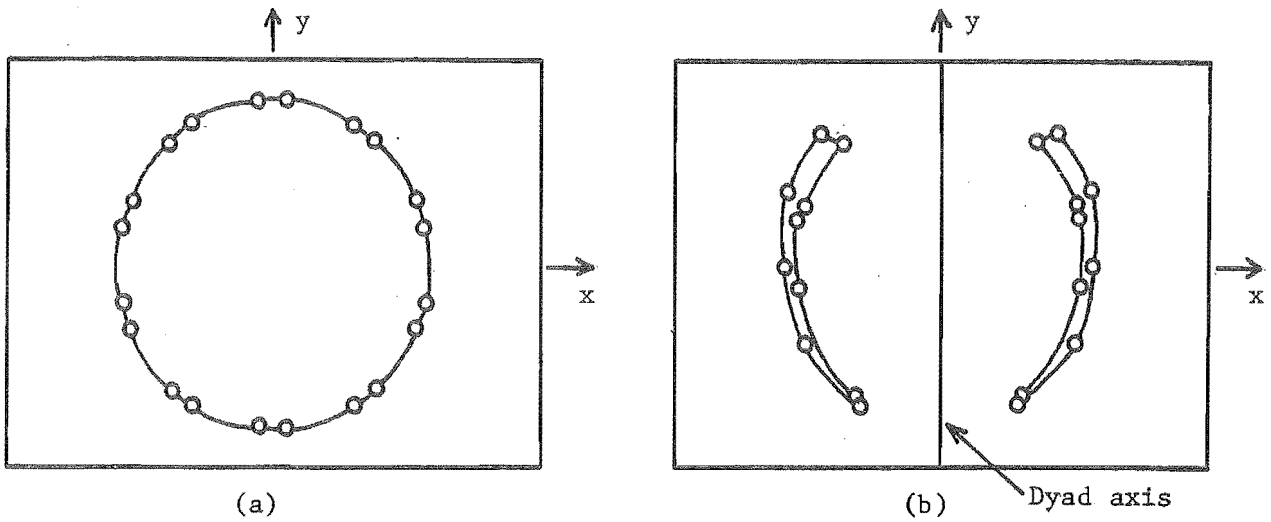


Fig. 2.5 Projections of phosphorus atoms onto the x-y plane.
(a) DH - Arnott and Hukins (1972a).
(b) SBS - Rodley *et al.* (1976).

where r is the radius from the centre of the phosphorus atom and r_o is its "effective radius". It was found that r_o = 0.75 Å gave the best fit to $\Phi_1(x,0)$ and $\Phi_1(0,y)$ in Fig. 2.4 in the region for which x and y are less than 2 Å, and this value of r_o was used throughout. Using (2.17) allows (2.14) and (2.15) to be evaluated analytically giving $\Phi_\ell(x,0)$ and $\Phi_\ell(0,y)$ in terms of the coordinates of the phosphorus atoms.

Fig. 2.6 shows $\Phi_1(x,0)$ and $\Phi_1(0,y)$ for the DH. The DH phosphorus coordinates used in this section and §2.3 are those given by Arnott and Hukins (1972a) and are illustrated in Fig. 2.5(a). The distance between the phosphorus atoms and most of the atoms in the sugars and the bases is greater than about 6 Å. This means that the Phi functions of models represented only by their phosphorus atoms can be reliably compared with those calculated from the data only for x and y less than about 6 Å. Inspection of Fig. 2.6 shows that $\Phi_1(x,0)$ and $\Phi_1(0,y)$ are both of the same general shape. This is to be expected as the symmetry of the DH implies that Pattersons in the x-y plane would be independent of direction. However the curves in Fig. 2.4 are quite different in the x and y directions, suggesting that the DNA molecule does not have the same symmetry as the DH.

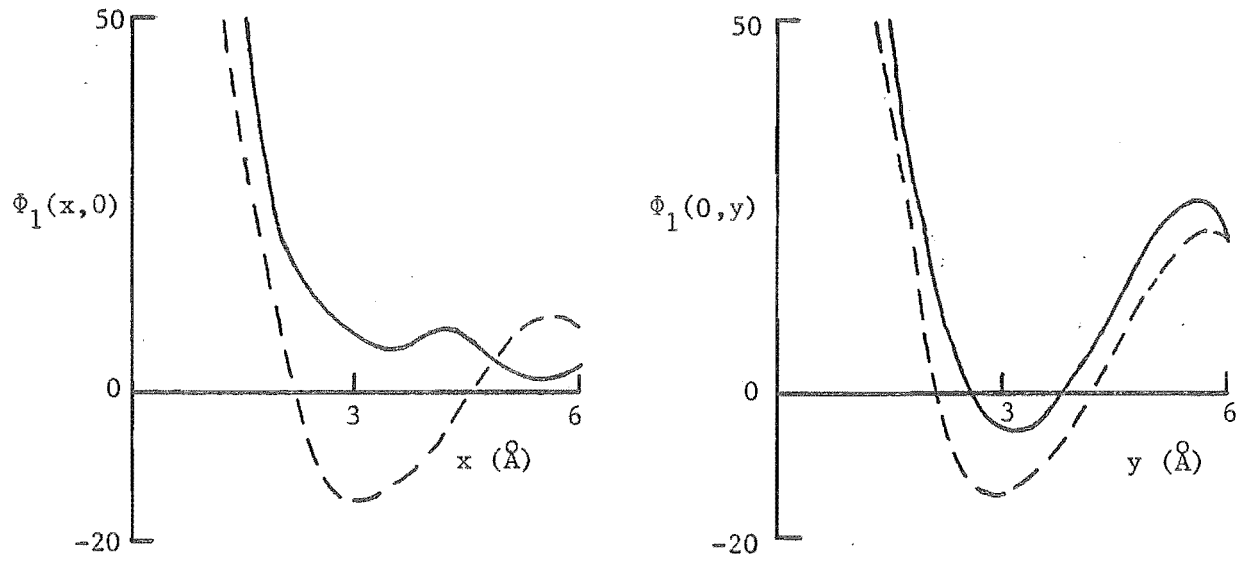


Fig. 2.6 The Phi functions for the data (——) and the DH model (---) given by Arnott and Hukins (1972a).

Fig. 2.7 shows the Phi functions calculated using the phosphorus coordinates for the SBS model given in Rodley *et al.* (1976). These coordinates are illustrated in Fig. 2.5(b). The SBS structure has an axis of symmetry as shown in Fig. 2.5(b) which is called a dyad axis.

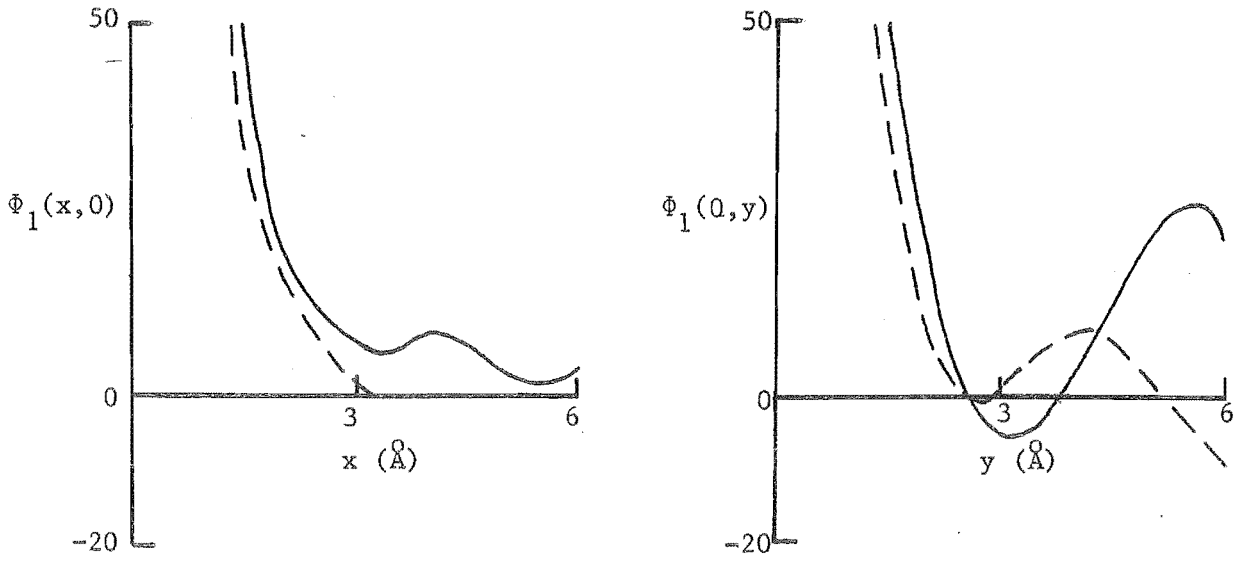


Fig. 2.7 The Phi functions for the data (——) and the SBS model (---) given by Rodley *et al.* (1976). $\theta = 0^{o}$.

The significance of the dyad axis is discussed in more detail in §2.4, and it suffices to say here that one half of the repeat unit can be generated by rotation of the other half by 180° about the dyad axis. When the molecule is viewed in projection, the dyad axis becomes a reflection axis of symmetry as shown in Fig. 2.5(b). The orientation of the SBS model in the unit cell is defined by the angle, denoted here by θ, between the dyad axis and the y-axis. The Phi functions for θ = 0° are shown in Fig. 2.7, and a comparison with those in Fig. 2.4 shows that they fit the data about as well as the DH.

The SBS coordinates given by Rodley *et al.* (1976) have a nett rotation (called a long range twist) of 35° per 10 base pairs (or 33.8 Å) because the right-handed sections rotate more than the left-handed sections. However the diffraction data indicates that the molecule has a repeat unit of 33.8 Å suggesting that it has no long range twist in the paracrystalline state. This twist may be present *in vivo* but absent, because of crystal packing forces, in paracrystalline samples. The SBS coordinates were therefore adjusted (minimally disturbing the molecule's stereochemistry) to remove this long range twist (see table 2.2(a)).

Table 2.2 SBS Phosphorus Coordinates (Å)

Phosphorus atom	(a) Untwisted (θ = 0°)			(b) Untwisted and modified (θ = 30°)		
	x	y	z	x	y	z
A6	-7.22	-4.56	-4.55	-3.97	-7.56	-4.55
A7	-3.39	-7.71	-7.65	0.92	-7.90	-7.65
A8	-3.46	-8.03	-11.50	-0.20	-8.68	-11.50
A9	-7.37	-3.44	-11.95	-4.66	-6.66	-11.95
A10	-7.90	0.32	-14.80	-7.00	-3.67	-14.80
B6	8.81	0.01	1.00	7.62	4.41	1.00
B7	8.38	4.56	-2.37	4.98	8.14	-2.37
B8	5.55	7.98	-5.67	0.28	9.69	-5.67
B9	4.06	7.77	-11.85	-0.37	8.76	-11.85
B10	7.16	4.56	-15.68	3.92	7.53	-15.68

The dyad axis lies in the x-y plane, passes through the origin and is inclined at an angle θ to the y-axis. Phosphorus atoms A1 - A5 and B1 - B5 can be generated from the above coordinates by a rotation of 180° about the dyad axis.

Since the Phi functions are forms of Pattersons, they can be interpreted in terms of interatomic vectors as long as the weighted averaging over z is taken into account. This means that small adjustments can be made to the atomic coordinates in order to change particular features of the Phi functions. Modifications were made to the SBS phosphorus coordinates to make the Phi functions match the data as closely as possible (see Table 2.2 (b)). It was also found that $\theta = 30^{o}$ gave the best agreement with the data. Only small changes were made to the phosphorus coordinates so that there was only minimal disturbance to the stereochemistry. In Fig. 2.8 the Phi functions of this modified SBS structure are compared with those of the DH and the data. It is apparent that the Phi functions for the SBS model follow the data more closely than do those for the DH.

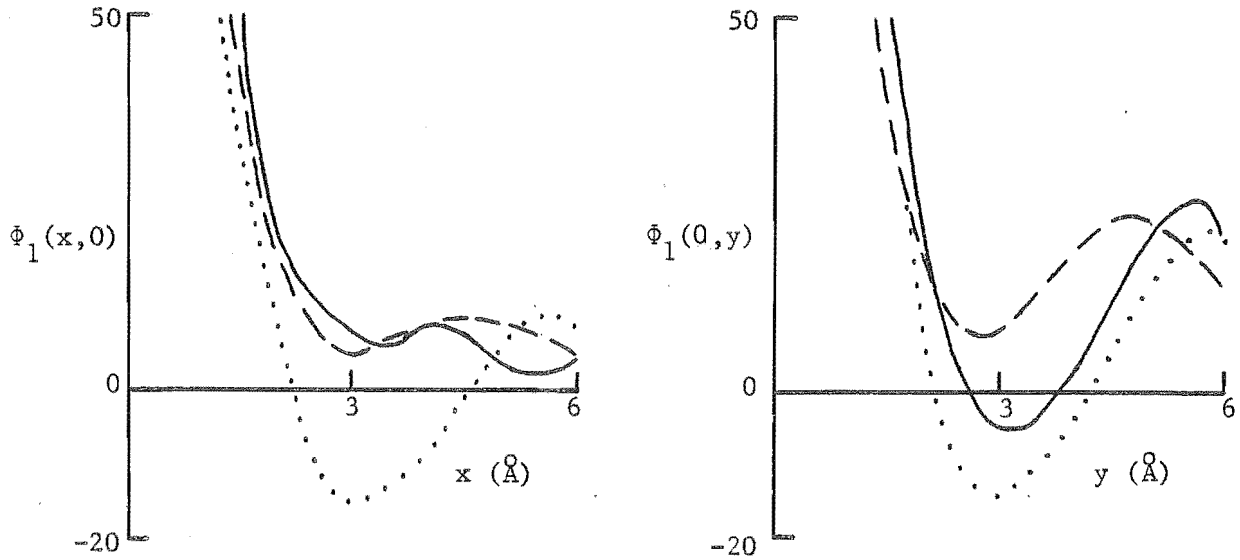


Fig. 2.8 The Phi functions for the data (——), the DH (···) and the modified SBS model (---) for $\theta = 30^{o}$ (see text).

These results are far from conclusive because of the simplified representations of the molecules used. However they do indicate a general consistency between the SBS model and the diffraction data, suggesting that this alternative structure is worthy of further study.

2.3 FIBRE ANALYSIS

Fibre samples differ from paracrystals because they are not composed of crystallites. The molecules are randomly rotated about their axes with respect to each other but all the molecular axes are parallel. Because a fibre sample is periodic in the z direction, the diffraction pattern exists on discrete planes (called layer planes), with $w = \ell/c$, in reciprocal space. Since there is no periodicity in the x-y plane, a continuous diffraction pattern is recorded on each layer plane. It is convenient to introduce cylindrical polar coordinates (r,θ,z) and (R,ψ,w) into real and reciprocal space respectively. Using (1.70), the diffraction pattern of a single molecule is given by

$$F_\ell(R,\psi) = \int_0^c \int_0^\infty \int_0^{2\pi} f(r;\theta,z) \exp(i2\pi(Rr\cos(\psi-\theta) + \ell z/c))\, r\, d\theta dr dz . \quad (2.18)$$

Since the individual molecules in the fibre sample are randomly rotated, what is measured is the angular average of the diffracted intensity of a single molecule. Hence the intensity $\Omega_\ell(R)$ of the fibre pattern on the ℓ^{th} layer line is given by

$$\Omega_\ell(R) = (1/2\pi) \int_0^{2\pi} |F_\ell(R,\psi)|^2 \, d\psi . \quad (2.19)$$

Because only the angular average of the diffraction pattern is available, it is not possible to calculate the Patterson $P(r;\theta,z)$ from the data $\Omega_\ell(R)$. However the cylindrically averaged Patterson $\tilde{P}(r,z)$, introduced in §2.2, can be computed. Transforming (2.7) to polar coordinates gives

$$\tilde{P}(r,z) = (1/2\pi) \int_0^{2\pi} P(r;\theta,z)\, d\theta . \quad (2.20)$$

The Patterson of a single molecule is obtained by Fourier transforming the diffraction pattern so that

$$P(r;\theta,z) = \sum_{\ell=-\infty}^{\infty} \int_0^\infty \int_0^{2\pi} |F_\ell(R,\psi)|^2 \exp(i2\pi(rR\cos(\theta-\psi) + \ell z/c))\, R\, d\psi dR . \quad (2.21)$$

Note that $P(r;\theta,z)$ takes the form of a Patterson in the z direction and an autocorrelation in the r-θ plane. Substituting from (2.21) into (2.20), using an integral representation of $J_0(x)$ (Watson, 1966, §2.2) and making use of (2.19) gives

$$\tilde{P}(r,z) = \sum_{\ell=-\infty}^{\infty} \exp(i2\pi\ell z/c) \int_0^{\infty} \Omega_\ell(R)\, J_o(2\pi rR)\, R\, dR \qquad (2.22)$$

showing that $\tilde{P}(r,z)$ can be calculated from the data $\Omega_\ell(R)$.

At this point (2.8) relating $\tilde{P}(r,z)$ to the paracrystalline diffraction pattern is conveniently derived. For a paracrystalline sample $\Omega_\ell(R)$ is the angular average of the $|F_{hk\ell}|^2$ so that

$$\Omega_\ell(R) = \sum_{h,k=-\infty}^{\infty} |F_{hk\ell}|^2\, \delta(R - (h^2/a^2 + k^2/b^2)^{\frac{1}{2}}) \ . \qquad (2.23)$$

Substituting (2.23) into (2.22) and making use of (2.1) gives (2.8).

Returning to the fibre patterns, the cylindrically averaged Patterson is given by

$$\tilde{P}(r,z) = \int_0^{2\pi} \int_{-\infty}^{\infty} f(x,y,z')\, f(x - r\cos\theta,\ y - r\sin\theta,\ z'-z)\, dxdydz'\, d\theta \ . \qquad (2.24)$$

Equation (2.24) is not particularly informative because the electron density is a function of Cartesian coordinates. An expression in terms of cylindrical polar coordinates is derived as follows. Substituting from (2.18) into (2.19) and reducing one of the angular integrations to a Bessel function gives

$$\Omega_\ell(R) = \int_0^{c} \int_0^{\infty} \int_0^{2\pi} Q(r,r',\theta,z)\, J_o(2\pi RB) \exp(i2\pi\ell z/c)\, rr'\, d\theta dr dr' dz \qquad (2.25)$$

where

$$Q(r,r',\theta,z) = \int_0^{c} \int_0^{2\pi} f(r;\phi,\zeta)\, f(r';\theta+\phi, z+\zeta)\, d\phi d\zeta \qquad (2.26)$$

and

$$B = (r^2 + r'^2 - 2rr'\cos\theta)^{\frac{1}{2}} \ . \qquad (2.27)$$

On substituting from (2.25) into (2.22), the integral over the Bessel functions and the summation both reduce to delta functions giving

$$\tilde{P}(\rho,z) = \int_0^{\infty} \int_0^{2\pi} Q(r,r',\theta,z)\, \delta(\rho - B)\, rr'\, d\theta dr dr'/\rho \ . \qquad (2.28)$$

The interpretation of (2.28) is rather complicated, however, if the electron density can be approximately represented as being confined to a circular cylinder of radius $\hat{r}$ so that

$$f(r;\theta,z) = f(\theta,z)\ \delta(r-\hat{r})\ , \tag{2.29}$$

then (2.28) simplifies to

$$\tilde{P}(r,z) = \int_{o}^{2\pi} Q(\theta,z)\ \delta(r - 2\hat{r}\ \sin(\theta/2))\ d\theta \tag{2.30}$$

where

$$Q(\theta,z) = \hat{r}^2 \int_{o}^{c}\int_{o}^{2\pi} f(\phi,\zeta)\ f(\theta+\phi,z+\zeta)\ d\phi\ d\zeta\ . \tag{2.31}$$

The θ dependence of $Q(\theta,z)$ has the form of an angular autocorrelation which is expected to be useful for comparing SBS and DH structures. The argument of the delta function in (2.30) is zero for two values of θ so that (2.30) may be written as

$$(\hat{r}/2)\ \cos(\theta/2)\ \tilde{P}(2\hat{r}\ \sin(\theta/2),z) = Q^{+}(\theta,z) \tag{2.32}$$

where $Q^{+}(\theta,z)$ is the even part of $Q(\theta,z)$ defined by

$$Q^{+}(\theta,z) = (Q(\theta,z) + Q(-\theta,z))/2 \tag{2.33}$$

and

$$\theta = 2\ \sin^{-1}(r/2\hat{r})\ . \tag{2.34}$$

$Q^{+}(\theta,z)$ is called here the <u>angular Patterson</u>. Equation (2.32) allows the fibre diffraction data and model structures to be compared, since the angular Patterson can be computed from the data and models using (2.22) and (2.31) respectively. Since the phosphorus atoms of both the SBS and DH models of DNA lie approximately on the surface of a cylinder, the angular Patterson can be applied to this representation of the models.

The angular Patterson is calculated from the fibre pattern using (2.22) in the form

$$\tilde{P}(r,z) = \sum_{\ell=o}^{L} \varepsilon_{\ell}\ \cos(2\pi\ell z/c) \int_{o}^{\hat{R}} \Omega_{\ell}(R)\ J_{o}(2\pi rR)\ R\ dR \tag{2.35}$$

where L and $\hat{R}$ are the maximum values of ℓ and R for which $\Omega_{\ell}(R)$ is recorded. The fibre diffraction data are taken from Feughelman *et al.* (1955) for which $L = 10$ and $\hat{R} = 0.026\ \text{Å}^{-1}$. Values of $\Omega_{o}(R)$ for $R < 0.05\ \text{Å}^{-1}$ are not available (and so have been set to zero) since they are obliterated by the incident X-ray beam. The values of $\Omega_{10}(R)$ for $R < 0.05\ \text{Å}^{-1}$ have been extrapolated from the data given by Feughelman *et al.* (1955). Fig. 2.9 shows the angular Patterson calculated from the data.

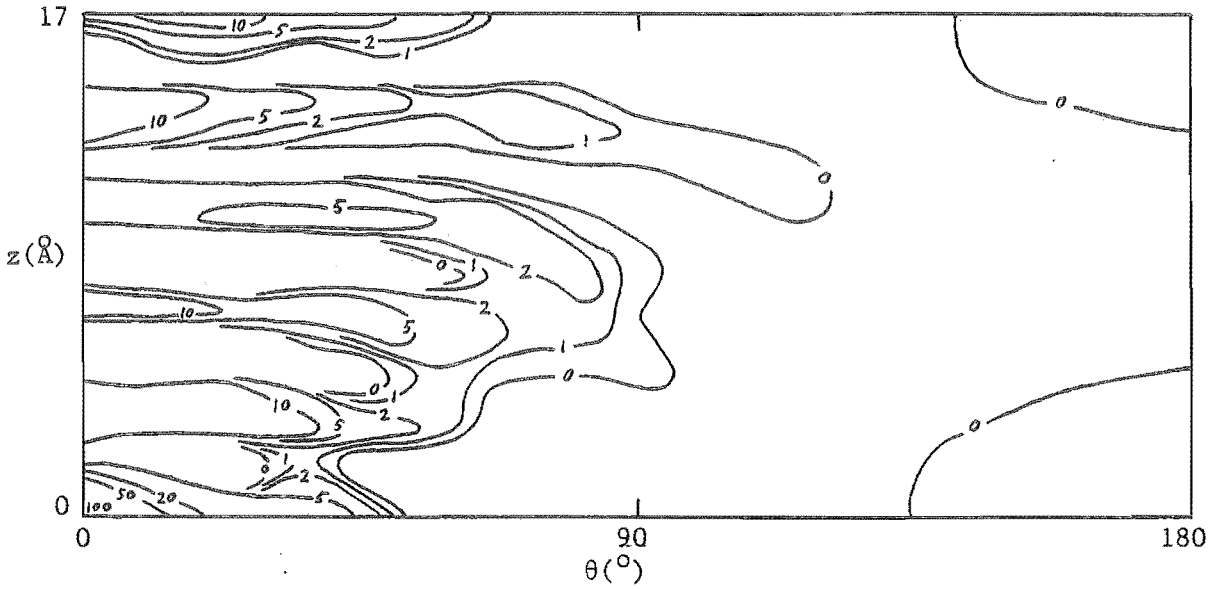


Fig. 2.9 $Q^{+}(\theta,z)$ calculated from the fibre diffraction data of Feughelman *et al.* (1955).

Note that the angular Patterson is symmetric about the line z = 17 Å so it only need be displayed on the interval 0 < z < 17 Å. It is also worth noting that the angular Patterson should be a positive function. The negative parts (not shown in Fig. 2.9) are due to $\Omega_{\ell}(R)$ only being measured in a finite region of reciprocal space - this also limits the resolution. Reference to (2.31) shows that peaks in Q^{+} correspond to interatomic "vectors" (in θ and z) in the structure. The peaks in the angular Patterson calculated from the data are shown in Fig. 2.10. The peaks spaced by 3.4 Å along the z-axis correspond to the base pairs which are stacked parallel to each other with this spacing along the molecular axis.

Since the angular Patterson is most easily interpreted in terms of its peaks, it is calculated from model structures whose atoms are represented as points (i.e. the electron density is represented as a delta function) in order to accentuate these peaks. Fig. 2.11 shows Q^{+} for a point representation of the phosphate groups of the DH. Note that Q^{+} peaks on straight lines in the r-θ plane which is a result of the phosphorus atoms lying on helices. Comparing Figs 2.9 and 2.11 shows that while the ridges in Fig. 2.9 tend towards the dotted lines, there is no evidence of the strong peaks present in Fig. 2.11.

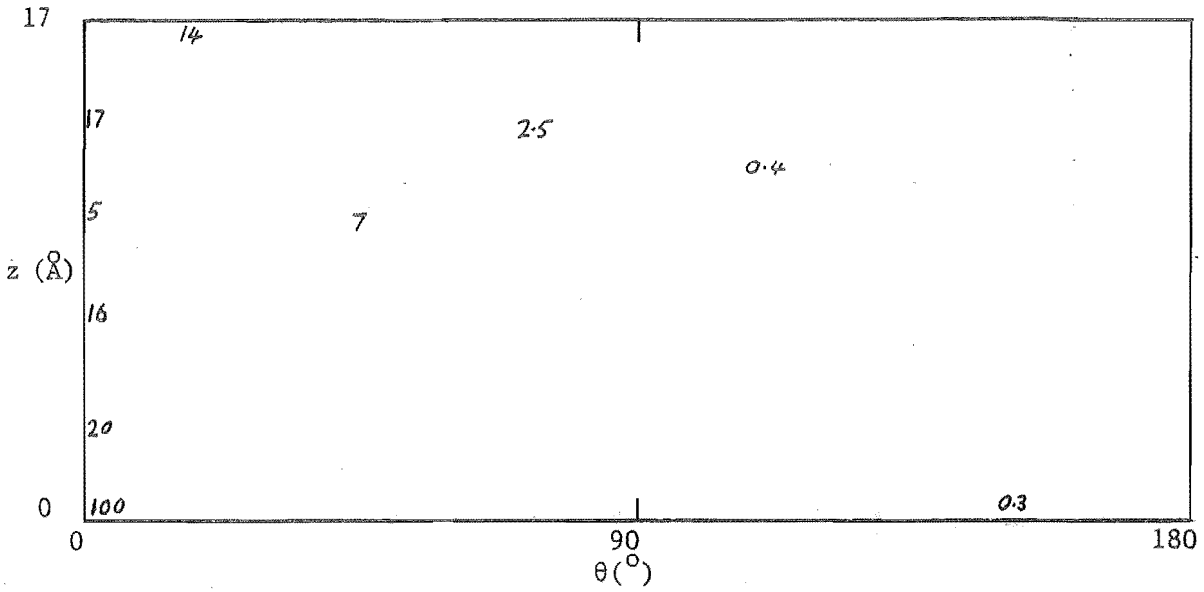


Fig. 2.10 Positions and amplitudes of the peaks in Fig. 2.9.

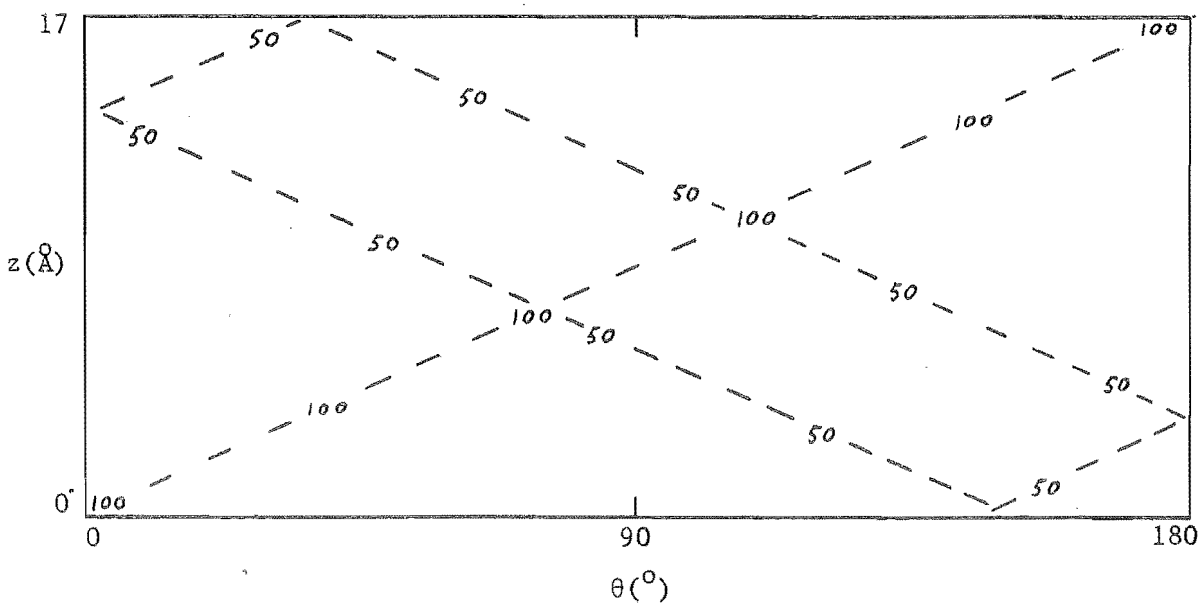


Fig. 2.11 $Q^{+}(\theta,z)$ for a point representation of the DH model.

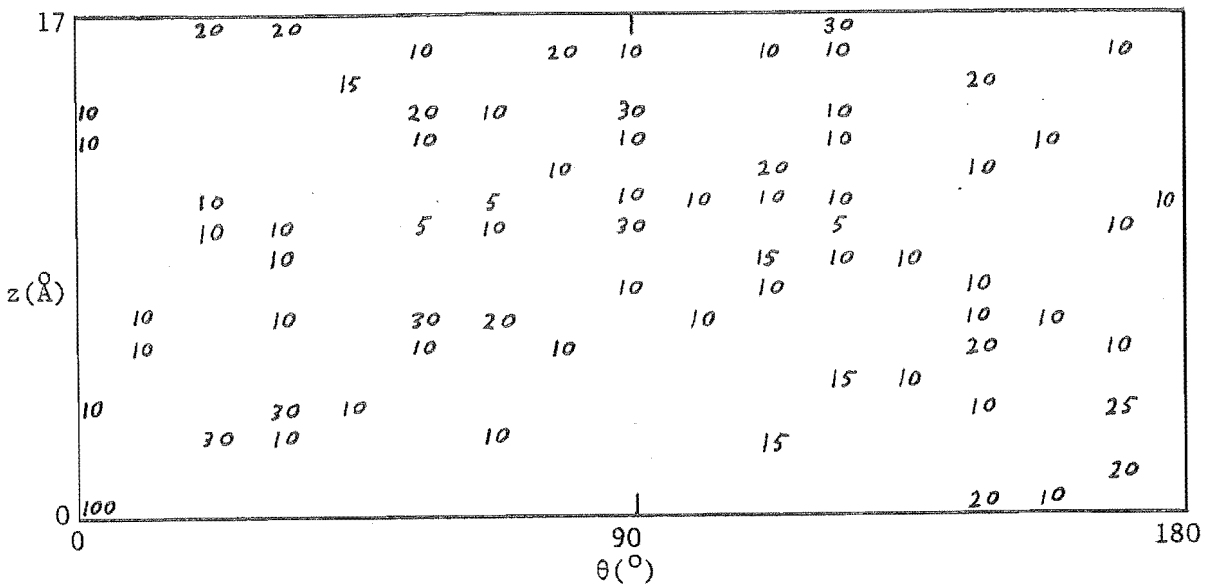


Fig. 2.12 $Q^+(\theta,z)$ for a point representation of the SBS model (see text).

Fig. 2.12 shows Q^+ for a point representation of the phosphorus atoms in the SBS model. The SBS coordinates used are those of Rodley *et al.* (1976) which have been untwisted as described in §2.2 (see table 2.2(a)). The peaks of Q^+ for the SBS model are not highly localised like the DH. However, Q^+ calculated from the data has six peaks on the z-axis ($\theta = 0$), and the DH and SBS models have two and four peaks respectively near the z-axis. Inspection of Figs 2.10 to 2.12 shows that the peaks of Q^+ near the z-axis for the SBS model, accord more closely with the data than do those of the DH.

The above results suggest that it is worth examining Q^+ along the z-axis. Making use of (2.32) shows that

$$Q^+(0,z) = (\hat{r}/2)\ \tilde{P}(0,z) \tag{2.36}$$

and so, referring to (2.24), the <u>axial Patterson</u> $P_{ax} = P_{ax}(z)$ may be defined by .

$$P_{ax}(z) = (1/\pi\hat{r})\ Q^+(0,z) = \iiint_{-\infty}^{\infty} f(x,y,\zeta)\ f(x,y,\zeta-z)\ dxdyd\zeta\ . \tag{2.37}$$

Inspection of (2.37) shows that P_{ax} takes the form of a Patterson along the z-axis. In order to take account of the peaks of Q^+ which are close to (but not on) the z-axis, the finite size of the atoms must be taken into account

when computing realistic axial Pattersons. For this reason P_{ax} is calculated from (2.37) using Gaussian atomic electron densities as described in §2.2. Axial Pattersons calculated in this manner for the DH and SBS models are shown in Fig. 2.13.

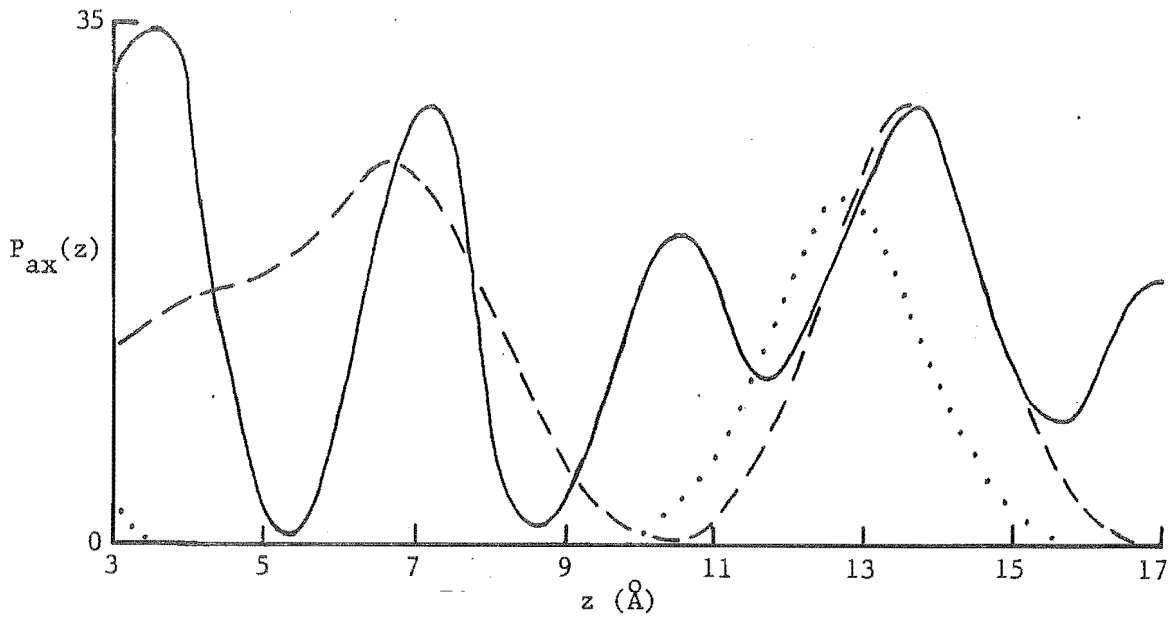


Fig. 2.13 Axial Pattersons calculated from the diffraction data (——), the DH model (•••) and the SBS model (---) (see text).

Making use of (2.35) to (2.37) shows that the axial Patterson is related to the data by

$$P_{ax}(z) = \sum_{\ell=o}^{L} \varepsilon_\ell \cos(2\pi\ell z/c) \int_o^{\hat{R}} \Omega_\ell(R)\, R\, dR\,. \qquad (2.38)$$

Inspection of (2.38) shows that $\Omega_o(R)$ does not affect the variation of $P_{ax}(z)$ with z but only adds a constant value. Since $\Omega_o(R)$ is not known for $R < 0.05\ \text{Å}^{-1}$, the value of this constant is unknown. Hence a constant value has been added to the axial Patterson calculated from the data (displayed in Fig. 2.13) and subsequently scaled so that it matches those calculated from the models as well as possible.

Since $P_{ax}(z)$ is a Patterson along the molecular axis, it is sensitive to the relative angular positioning of the nucleotides. This is because nucleotides with similar angular positions produce large values of $P_{ax}(z)$ at values of z equal to their axial spacing. The peaks spaced by

approximately 3.4 Å in $P_{ax}(z)$ calculated from the data are due to overlap of the nucleotides (particularly the base pairs) which are spaced by about 3.4 Å. As described in §2.2, an accurate measure of the angular positioning of the nucleotides would allow one to decide between the SBS and DH models. Hence the axial Patterson is considered to be particularly useful for comparing these two models with the data.

Inspection of Fig. 2.13 shows that the axial Patterson for the SBS model fits the data better than the DH, particularly because the latter does not produce a peak at about 6.8 Å. The nucleotide pairs on each side of the bend region (where the molecule changes its sense of rotation) of the SBS structure have similar angular positions and are spaced by about 6.8 Å which explains the presence of this peak. Because the DH does not change its sense of rotation, nucleotide pairs separated by 6.8 Å have quite different angular positions and so produce little amplitude at this axial distance. Hence the axial Pattersons displayed in Fig. 2.13 appear to provide quite strong evidence in favour of the SBS model over the DH. The presence of the base pairs has a strong influence on the axial Patterson since because they are planar and perpendicular to the molecular axis, they overlap significantly as they shift along the molecular axis. However, as described above, it is difficult to see how including the effects of the bases would produce significant amplitude in the axial Patterson at z = 6.8 Å. However before these results could be taken too seriously, the axial Patterson needed to be calculated using realistic representations of all the atoms in the molecule. This is discussed in §2.5.

Note that in moving from the angular Patterson $Q^{+}(\theta,z)$ to the axial Patterson $P_{ax}(z)$, the restriction that the electron density must be confined to the surface of a cylinder has been removed. This can be seen from (2.37) and it allows the axial Patterson to be calculated for a more realistic representation of the molecule. It is worth noting that Bates *et al.* (1980) conclude that, in some circumstances, it may be safer to base structural refinement on fibre, rather than paracrystalline, diffraction data.

2.4 STEREOCHEMICAL REFINEMENT

The stereochemistry of a molecule refers to the geometrical relationships between constituent atoms and groups of atoms. Some of these geometric parameters (such as bond lengths and bond angles) have been found to vary

very little over a large number of different molecules. Others (such as some conformation angles) are known to vary over much larger ranges of values (Arnott, 1970). Nonetheless, Hingerty (1979) has shown that even quite small changes in some of these parameters produce large changes in helical polynucleotide conformations. The ranges of values which these parameters can adopt are referred to as the stereochemical constraints on the molecule.

The atomic coordinates presented by Rodley *et al.* (1976) demonstrate the feasibility of the SBS structure but do not conform accurately to these stereochemical constraints. Following the supportive evidence of the X-ray calculations described in §2.2 and §2.3 and those of McKinnon (1980) (reported in Bates *et al.*, 1980), using more accurate representations of the molecule, it was considered essential to examine the stereochemical viability of the SBS structure.

As described in §2.1, the DNA molecule is made up of two strands of nucleotides, each consisting of a phosphate, a sugar and one of four bases. The bases are planar groups of known fixed conformations and the primary concern here is with the sugar-phosphate backbone. The sugar phosphate part of a nucleotide is called a nucleoside, the constituent atoms of which are shown in Fig. 2.14.

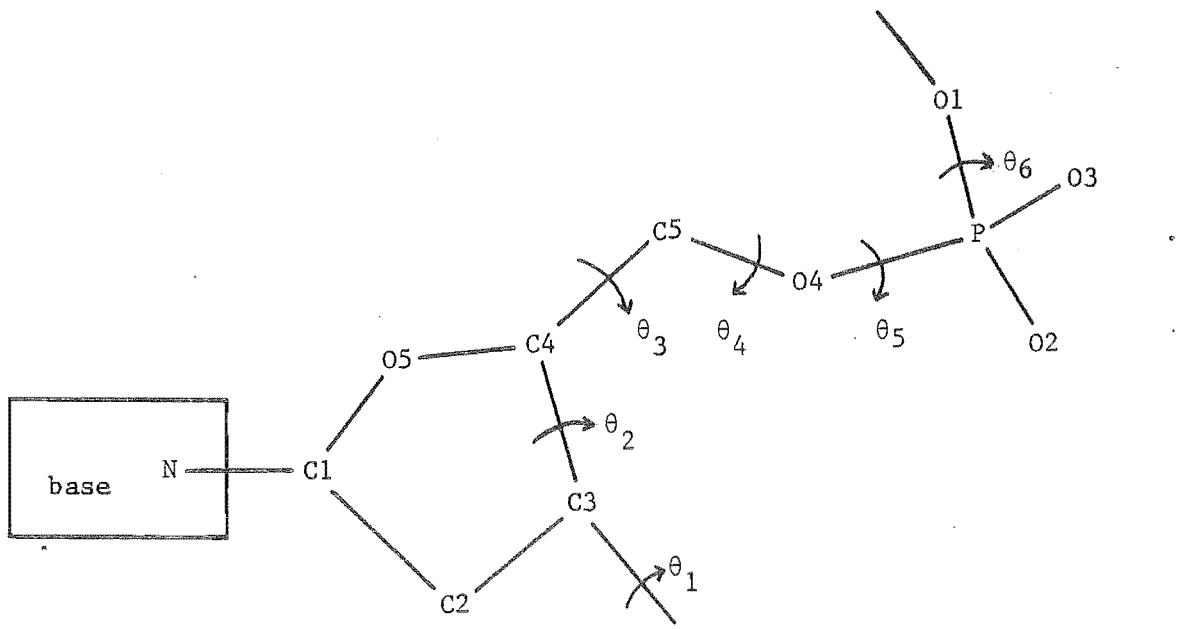


Fig. 2.14 Constituent atoms of a nucleoside.
$\theta_1 - \theta_6$ = conformation angles.

The conformation of the phosphate group is fixed. The sugar ring can adopt a number of conformations which are discussed later. Since the bond lengths (distance between two bonded atoms) and bond angles (angle between two adjacent bonds) are fixed, a single nucleoside has six degrees of freedom due to rotation about six bonds as shown in Fig. 2.14. These rotations are described by six conformation angles (Arnott, 1970) denoted by θ_1 to θ_6 in Fig. 2.14. Polynucleotide structures have been observed with a wide range of conformation angles (Arnott and Chandrasekaran, 1981). Since the DH is a regular structure in which successive nucleotides are shifted and rotated by 3.38 Å and 36° respectively, the backbone is completely described by these six conformation angles and the sugar conformation. The SBS model does not possess this high degree of symmetry as the conformation of each nucleotide is different, especially at the bend regions where the sense of rotation changes.

In this section the symmetry of the SBS structure, the physical model from which initial atomic coordinates were obtained, the stereochemical parameters considered and the subsequent refinement procedure are described.

The original SBS structure (Rodley *et al.*, 1976) contains a long range twist of about 35° per ten base pairs. It appears from the subsequent stereochemical investigation described here that this is an essential feature of the structure. This means that a ten base pair unit does not form a repeat unit. However successive ten base pair units can be generated by an axial shift (by a distance denoted by c), followed by a rotation about the molecular axis by the long range twist angle denoted here by θ. Since the two chains are antiparallel, one can be generated from the other by a rotation of 180° about an appropriately positioned dyad or two-fold axis which is perpendicular to and passes through the molecular axis. Rotation about the dyad axis allows a five base pair unit to be generated from an adjacent five base pair unit. Let $T(\theta,c)$ represent a translation of distance c along the molecular axis followed by a rotation by θ about the same axis. Let $R(\phi,\zeta)$ represent a rotation by 180° about a line which is perpendicular to the molecular axis with angular coordinate ϕ and axial coordinate ζ. It is easily shown that

$$T(\theta,c) = R(\phi+\theta/2,\zeta+c/2)\ R(\phi,\zeta)\ . \qquad (2.39)$$

Equation (2.39) shows that the translation and axial rotational symmetry is equivalent to two rotations about two dyad axes spaced by c/2 and rotated with respect to each other by $\theta/2$. Hence the SBS structure is completely

described by the coordinates of a five base pair unit together with these two dyad axes.

In order to study the stereochemical viability of the SBS structure it was necessary to construct an accurate physical model to obtain the initial atomic coordinates. The model pertains to the fibre form of DNA since the long range twist is retained and so only the intramolecular (and not the intermolecular) stereochemistry need be considered. In paracrystalline samples, crystal packing forces could reduce the long range twisting to zero. The base pairs in the original model (Rodley *et al.*, 1976) were constrained such that they were perpendicular to, and their midpoints coincided with, the molecular axis. This produced unacceptably short distances between non-bonded atoms in the bend regions. It was found that by relaxing these constraints, it is possible to build a stereochemically acceptable model.

The physical model was constructed by Dr. Rodley of the Chemistry department using "Kendrew (2 cm = 1 Å) molecular units" for the sugar-phosphate backbone (except the hydrogen atoms) and perspex plates to represent the base pairs. The model consists of seven nucleotide pairs in order to allow the positions of the two two-fold axes at each end of the central five base pair unit to be estimated. The central five nucleotide pairs (with the bend region in the centre) are numbered 6 to 10. The two two-fold axes (labelled L_2 and R_2) lie in the centres of the left-handed and right-handed regions between dinucleotide pairs 5 and 6 and between 10 and 11 respectively. The model was adjusted to satisfy the stereochemical constraints as well as possible. The coordinates of the backbone atoms and the glycosidic nitrogen atoms (the base atoms bonded to the sugars - see Fig. 2.14) were measured (by Miss G.F. Rodley) using sightings from two theodolites. The theodolite measurements had an accuracy equivalent to about 0.002 Å for each Cartesian axis for the atomic coordinates. This is well within the stereochemical tolerance required and so the limiting factor was the accuracy of building the physical model. The coordinates were fed into the computer where the refinement procedure was carried out.

The standard values for the backbone bond lengths (rounded to 0.01 Å) and bond angles are taken from Arnott and Hukins (1972b) (sugars) and Arnott (1970) (phosphate) and are listed in table 2.3. These values have been determined by high resolution X-ray crystallographic analysis of monomers (single nucleotides or nucleosides) related to nucleic acids.

Table 2.3 Standard bond lengths (Å) and bond angles (°).

N - C1	1.49,1.47*	C4 - C5	1.52
C1 - C2	1.53	C5 - O4	1.44
C2 - C3	1.53	O4 - P	1.60
C3 - C4	1.53	P - O1	1.60
C1 - O5	1.42	P - O2,O3	1.48
O5 - C4	1.46	O1 - C3	1.42
N - C1 - O5	108.0	C4 - C5 - O4	110.0
N - C1 - C2	113.7	C5 - O4 - P	119.0
O5 - C1 - C2	107.6,105.6*	O4 - P - O1	101.4
C1 - C2 - C3	101.4	P - O1 - C3	119.0
C2 - C3 - C4	102.5	O4 - P - O2	107.8
C3 - C4 - O5	104.0,105.5*,103.9**	O1 - P - O3	107.7
C1 - O5 - C4	109.7	O2 - P - O3	116.0
O5 - C4 - C5	108.9	O1 - C3 - C4	111.9,108.9*
C3 - C4 - C5	115.8	O1 - C3 - C2	113.3,109.3*,**

Values listed are for C3-endo. Those that are different for C2-endo and C3-exo are indicated by * and ** respectively.

The SBS coordinates were refined so that the bond lengths and bond angles are within ±0.025 Å and ±2.0° respectively of these standard values. These tolerances are close to the standard deviations of these parameters observed in various nucleotide structures (Arnott and Hukins, 1972b).

The distances between pairs of non-bonded atoms (excluding atoms of a pair which are bonded to a common atom) are referred to as contacts and must be greater than prescribed minimum values. Minimum allowable intramolecular contacts within macromolecular structures tend to be shorter than those between different molecules (Haschemeyer and Rich, 1967). The guide for allowable contacts is taken from the sum of the intramolecular nucleoside Van der Waals radii given by Haschemeyer and Rich (1967) which are listed in table 2.4. Short non-bonded distances to the nitrogen atoms do not occur in the SBS structure.

The five atoms making up the sugar ring are non-planar and this is referred to as the sugar pucker. The pucker can vary considerably with negligible changes occurring in the ring bond lengths and bond angles. The sugar pucker can be described by three conformation angles (Arnott, 1970)

Table 2.4 Minimum Intramolecular Contacts (Å).

Atom pairs	From Haschemeyer and Rich (1967)		Limit used in SBS refinement
	Normal limit	Rare limit	
C - O	2.85	2.65	2.75
C - C	3.00	2.80	2.80
O - O	2.70	2.50	2.75

or equivalently by the distances (denoted by d_2 and d_3) of the atoms C2 and C3 from the C1-O5-C4 plane (Arnott and Hukins, 1972b). The latter description is used here as the geometrical significance is more immediately apparent. The sign of d_2 (d_3) is positive if C2 (C3) is on the same side of the C1-O5-C4 plane as C5, and negative if it is on the other side. The terms C2-endo, C3-endo, C2-exo and C3-exo are commonly used as qualitative descriptions of sugar pucker. The prefix refers to the atom most distant from the C1-O5-C4 plane, and endo or exo signifies that this atom is on the same or the opposite side respectively of the C1-O5-C4 plane as C5. A survey of sugar puckers carried out by Arnott and Hukins (1972b) shows a wide variability in the values of d_2 and d_3 so it is considered appropriate to use sugar puckers which are reasonably close to one of those observed by them.

The bases are positioned on the backbone using the glycosidic bond (the N-C1 bond in Fig. 2.14) and then twisted and tilted into position so as to satisfy the stereochemical constraints. The bases are fixed planar groups and the coordinates are computed from the conformations given by Arnott and Hukins (1973). Each base can rotate about its glycosidic bond. This is referred to here as twist and is measured by the twist angle α. Zero twist ($\alpha = 0$) occurs when a line in the base plane which is perpendicular to the glycosidic bond is perpendicular to the molecular axis. The base can also be tilted so that the glycosidic bond does not lie in the base plane (Haschemeyer and Rich, 1967). The angle β between the glycosidic bond and the base plane is called the tilt angle. Haschemeyer and Rich (1967) have observed tilt angles up to 12.4°. These definitions of twist and tilt are different from those of Arnott (1970) which apply to regular helical structures. The angle between the normals to the two base planes of a base pair is called the dihedral.

The refinement procedure has three objectives:

1. Bring the bond lengths, bond angles, contacts and sugar puckers within the limits described above.

2. Obtain the positions of the dyad axes L_2 and R_2 so that connectivity between successive five base pair units is achieved with the same stereochemical constraints.

3. Position the bases on the backbone satisfying the same stereochemical constraints and align the two bases of each base pair as well as possible.

The stereochemical tolerances used are listed in Table 2.7.

The large number of complex interrelationships between all the stereochemical parameters described above would make the complete automation of the refinement procedure a computationally horrendous task. So in order to make the refinement tractable and obtain a sensible solution, it is necessary to carry it out in an interactive manner. The computer is used to calculate all the stereochemical parameters (described above) necessary to assist the operator in making appropriate adjustments to the atomic coordinates. The strategy of this refinement procedure is as follows.

1. Adjust the atomic coordinates until the bond lengths, bond angles, contacts and sugar puckers are within the required limits.

2. The dyad axis L_2 is initially estimated as being the line joining the midpoints of A5N-B6N and B5N-A6N (and similarly for R_2) where the coordinates of nucleotides 5 and 11 are obtained from the model.

3. Using these dyad axes, coordinates are generated for nucleotide pairs 5 and 11. The stereochemistry across the dyad axis is adjusted by moving atoms O1 and C3 and deriving new dyad axes. This step is repeated until the stereochemistry across the two-fold axes is within the required limits.

4. Each of the four bases (A, T, G or C) is located on each nucleoside and the position of the glycosidic nitrogens adjusted to minimise contacts to the backbone and improve the hydrogen bond lengths.

5. The bases are then tilted and twisted to minimise the remaining contacts and dihedral angles, and make the bases as coplanar as possible.

The resulting refined atomic coordinates with GC base pairs are given in table 2.8 and the stereochemical tolerances achieved are summarised in table 2.7. Other base pair arrangements can be accommodated by the same backbone structure. The contacts are all greater than 2.75 Å, the C-C contacts are all greater than 2.80 Å, and these are related to the Haschemeyer and Rich (1967) limits in table 2.4. The angle between the dyad axes L_2 and R_2 is 23.4°, they are spaced by 16.998 Å and their positions are given in table 2.8. This gives a long range twist of θ = 46.8° per ten dinucleotides and c = 33.996 Å. The sugar atom displacements d_2 and d_3 are listed in table 2.5 and each pair (d_2, d_3) is within 0.12 Å of that for one of the observed structures reported in table 1 of Arnott and Hukins (1972b).

Table 2.5 Sugar puckers.

Nucleoside	Pucker	d_2(Å)	d_3(Å)	Nucleoside	Pucker	d_2(Å)	d_3(Å)
A6	C3-endo	0.34	0.79	B6	C3-endo	0.10	0.74
A7	C3-endo	0.28	0.78	B7	C3-endo	0.20	0.76
A8	C3-endo	-0.28	0.43	B8	C3-endo	0.60	0.83
A9	C3-exo	0.10	-0.31	B9	C3-endo	0.59	0.77
A10	C3-endo	0.60	0.73	B10	C2-endo	0.39	-0.05

Base tilt angles are listed in table 2.6 and are all less than the maximum value of 12.4° observed by Haschemeyer and Rich (1967). The base twist and dihedral angles (see table 2.6) are all small, being less than 1° and 3° respectively. All hydrogen bond lengths are within the range 2.48 - 3.14 Å.

Table 2.6 Base pair geometry.

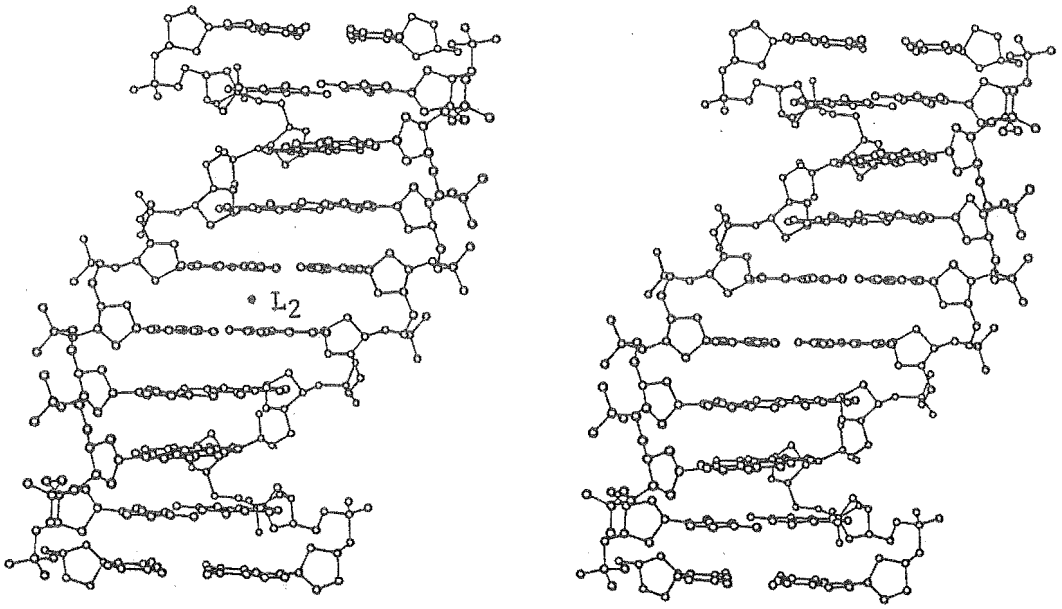
Base	Tilt, β (°)	Dihedral (°)
A6	-2.1	0.3
A7	0	2.7
A8	0	1.1
A9	-0.5 to 5.0	0.1
A10	0	1.2
B6	0	
B7	5.0	
B8	5.6	
B9	-3.9 to -10.4	
B10	6.0	

Tilt angles at A9 and B9 depend on the base.

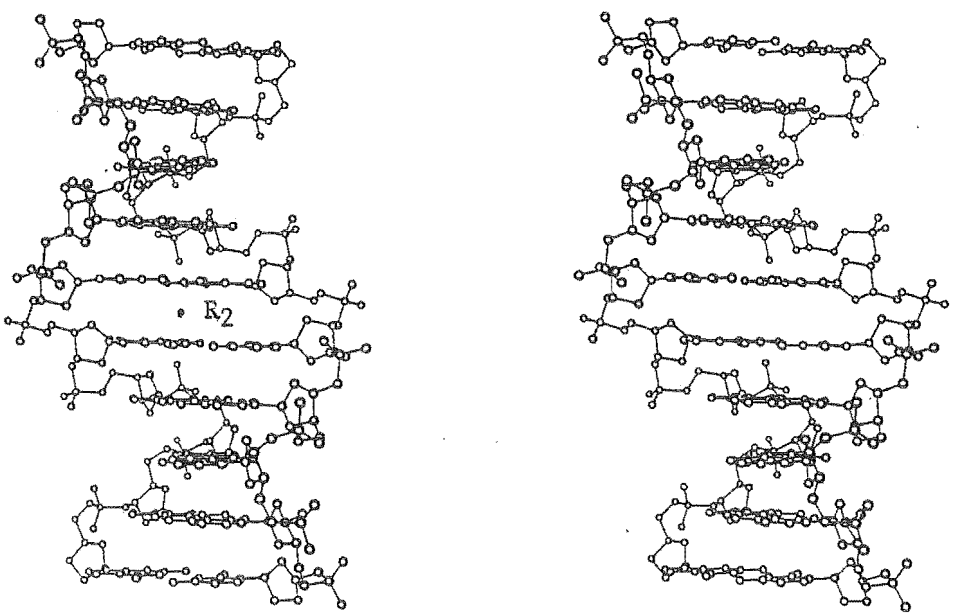
Table 2.7 Summary of stereochemical tolerances.

Parameter	Tolerance	Reference
Bond lengths	±0.02Å from standard values	Table 2.3
Bond angles	±2.0° from standard values	Table 2.3
Contacts	All <2.75 Å, C-C <2.80 Å	Table 2.4
Sugar puckers	d_2 and d_3 within ±0.12Å of observed values	Table 2.5
Base tilt (β)	All <11°	Table 2.6
Base twist (α)	All <1°	
Base pair dihedral	All <3°	Table 2.6
Hydrogen bond lengths	2.48 - 3.14 Å	
Long range twist (θ)	46.8° per 10 dunucleotides	

The refined structure outlined here demonstrates the stereochemical viability of the SBS model for DNA. Two stereodiagrams of the model viewed along L_2 and R_2 (see Fig. 2.15) show that it is a zig-zag structure since the centres of the base pairs do not lie above each other.



(a)



(b)

Fig. 2.15 Stereodiagrams of 10 base pair sections of the stereochemically refined SBS model.
(a) View along L_2 with the left-handed region in the centre.
(b) View along R_2 with the right-handed region in the centre.

Table 2.8 Backbone atomic coordinates (in Å) for the stereochemically refined SBS model.

Atom	x	y	z	Atom	x	y	z
A6C1	4.137	-1.956	-1.720	B6C1	-6.419	-3.851	-1.810
A6C2	5.470	-2.004	-0.929	B6C2	-7.542	-4.186	-2.810
A6C3	6.528	-1.876	-2.020	B6C3	-8.564	-4.864	-1.875
A6C4	5.812	-1.060	-3.075	B6C4	-8.431	-3.976	-0.620
A6C5	6.418	-1.160	-4.455	B6C5	-8.940	-4.598	0.685
A6O5	4.453	-1.559	-3.070	B6O5	-6.983	-3.792	-0.512
A6O4	5.851	-0.176	-5.375	B6O4	-8.562	-3.728	1.800
A6P	5.838	1.395	-5.007	B6P	-9.090	-2.193	1.865
A6O1	4.258	1.772	-5.035	B6O1	-7.745	-1.323	1.605
A6O2	6.539	2.148	-6.072	B6O2	-9.597	-1.909	3.277
A6O3	6.360	1.576	-3.635	B6O3	-10.042	-1.958	0.757
A7C1	1.099	-0.128	-5.015	B7C1	-7.606	-6.699	-5.365
A7C2	2.385	0.337	-4.285	B7C2	-8.371	-7.535	-6.415
A7C3	3.278	0.795	-5.443	B7C3	-9.126	-8.484	-5.495
A7C4	2.257	1.334	-6.465	B7C4	-9.467	-7.619	-4.300
A7C5	2.714	1.497	-7.905	B7C5	-9.849	-8.333	-3.000
A7O5	1.150	0.367	-6.370	B7O5	-8.234	-6.870	-4.095
A7O4	1.663	2.059	-8.685	B7O4	-9.859	-7.367	-1.930
A7P	1.097	3.535	-8.370	B7P	-10.892	-6.131	-2.060
A7O1	-0.361	3.453	-9.080	B7O1	-9.912	-4.877	-2.380
A7O2	1.934	4.546	-9.054	B7O2	-11.559	-5.912	-0.757
A7O3	0.936	3.706	-6.909	B7O3	-11.787	-6.362	-3.216
A8C1	-2.764	0.702	-8.655	B8C1	-7.160	-9.282	-8.840
A8C2	-1.893	1.797	-8.025	B8C2	-7.502	-10.105	-10.065
A8C3	-1.042	2.199	-9.222	B8C3	-8.293	-11.277	-9.495
A8C4	-2.100	2.235	-10.330	B8C4	-8.691	-10.831	-8.110
A8C5	-1.591	2.147	-11.750	B8C5	-8.821	-11.938	-7.065
A8O5	-2.989	1.109	-10.020	B8O5	-7.618	-9.970	-7.700
A8O4	-0.207	1.683	-11.792	B8O4	-9.275	-11.352	-5.815
A8P	0.897	2.637	-12.425	B8P	-10.672	-10.595	-5.725
A8O1	2.238	2.253	-11.605	B8O1	-10.288	-9.100	-6.070
A8O2	0.533	4.044	-12.144	B8O2	-11.176	-10.658	-4.335
A8O3	1.074	2.285	-13.852	B8O3	-11.582	-11.116	-6.770
A9C1	1.104	-0.717	-12.085	B9C1	-8.256	-6.289	-11.770
A9C2	1.914	-0.085	-10.955	B9C2	-9.070	-6.524	-13.050
A9C3	2.799	0.927	-11.679	B9C3	-10.365	-7.209	-12.530
A9C4	2.917	0.385	-13.100	B9C4	-10.320	-7.016	-11.030
A9C5	4.177	-0.462	-13.385	B9C5	-10.992	-8.116	-10.185
A9O5	1.768	-0.455	-13.297	B9O5	-8.917	-6.901	-10.685
A9O4	4.539	-1.248	-12.205	B9O4	-10.126	-9.282	-10.014
A9P	5.885	-2.157	-12.265	B9P	-10.593	-10.682	-10.695
A9O1	5.950	-2.603	-13.820	B9O1	-9.402	-11.708	-10.295
A9O2	5.704	-3.359	-11.421	B9O2	-10.622	-10.522	-12.166
A9O3	7.060	-1.314	-11.950	B9O3	-11.854	-11.133	-10.066
A10C1	2.999	-3.770	-15.425	B10C1	-7.876	-2.741	-15.200
A10C2	3.626	-3.523	-14.045	B10C2	-8.601	-2.684	-16.570
A10C3	5.141	-3.699	-14.303	B10C3	-9.962	-2.113	-16.225
A10C4	5.263	-3.813	-15.840	B10C4	-10.178	-2.509	-14.775
A10C5	6.341	-4.759	-16.298	B10C5	-11.172	-3.595	-14.493
A10O5	3.978	-4.247	-16.308	B10O5	-8.891	-2.845	-14.225
A10O4	6.035	-6.086	-15.880	B10O4	-10.629	-4.909	-14.760
A10P	6.924	-7.308	-16.343	B10P	-11.589	-6.214	-14.650
A10O1	6.229	-7.673	-17.740	B10O1	-11.543	-6.579	-13.069
A10O2	6.760	-8.428	-15.390	B10O2	-10.991	-7.325	-15.424
A10O3	8.312	-6.850	-16.573	B10O3	-12.971	-5.835	-15.021

Table 2.8 (cont.) Base pair (GC) coordinates

Atom	x	y	z	Atom	x	y	z
A6N9	3.489	-3.284	-1.785	B6N1	-5.348	-4.870	-1.815
A6C8	4.068	-4.536	-1.790	B6C6	-5.638	-6.199	-1.821
A6N7	3.206	-5.522	-1.799	B6C5	-4.652	-7.134	-1.826
A6C5	1.972	-4.883	-1.802	B6C4	-3.306	-6.654	-1.824
A6C4	2.135	-3.513	-1.793	B6N4	-2.298	-7.507	-1.828
A6C6	0.650	-5.423	-1.811	B6N3	-3.029	-5.342	-1.817
A6O6	0.303	-6.603	-1.820	B6C2	-4.033	-4.424	-1.813
A6N1	-0.319	-4.413	-1.810	B6O2	-3.818	-3.203	-1.807
A6O2	-0.058	-3.059	-1.801	B7N1	-6.157	-7.075	-5.270
A6N2	-1.130	-2.254	-1.801	B7C6	-5.767	-8.377	-5.305
A6N3	1.170	-2.549	-1.792	B7C5	-4.451	-8.717	-5.338
A7N9	1.014	-1.629	-5.086	B7C4	-3.507	-7.643	-5.333
A7C8	2.032	-2.560	-5.109	B7N4	-2.212	-7.900	-5.363
A7N7	1.618	-3.801	-5.175	B7N3	-3.901	-6.362	-5.298
A7C5	0.233	-3.687	-5.198	B7C2	-5.224	-6.047	-5.266
A7C4	-0.146	-2.361	-5.144	B7O2	-5.629	-4.875	-5.234
A7C6	-0.777	-4.694	-5.265	B8N1	-5.697	-8.938	-8.667
A7O6	-0.641	-5.916	-5.319	B8C6	-4.745	-9.909	-8.660
A7N1	-2.061	-4.137	-5.265	B8C5	-3.421	-9.596	-8.637
A7O2	-2.344	-2.788	-5.208	B8C4	-3.087	-8.206	-8.622
A7N2	-3.644	-2.459	-5.219	B8N4	-1.821	-7.831	-8.600
A7N3	-1.408	-1.844	-5.145	B8N3	-4.033	-7.257	-8.630
A8N9	-2.096	-0.653	-8.635	B8C2	-5.351	-7.594	-8.653
A8C8	-0.752	-0.963	-8.620	B8O2	-6.255	-6.746	-8.660
A8N7	-0.501	-2.248	-8.604	B9N1	-6.866	-6.819	-11.885
A8C5	-1.763	-2.832	-8.609	B9C6	-6.621	-8.156	-11.884
A8C4	-2.745	-1.864	-8.628	B9C5	-5.350	-8.639	-11.894
A8C6	-2.145	-4.208	-8.599	B9C4	-4.294	-7.676	-11.905
A8O6	-1.425	-5.205	-8.582	B9N4	-3.035	-8.074	-11.916
A8N1	-3.537	-4.355	-8.610	B9N3	-4.545	-6.359	-11.906
A8O2	-4.448	-3.319	-8.629	B9C2	-5.826	-5.900	-11.896
A8N2	-5.741	-3.673	-8.637	B9O2	-6.099	-4.690	-11.897
A8N3	-4.098	-2.036	-8.639	B10N1	-7.046	-3.977	-15.065
A9N9	0.863	-2.171	-11.955	B10C6	-7.618	-5.211	-15.054
A9C8	1.775	-3.206	-11.960	B10C5	-6.858	-6.338	-15.073
A9N7	1.229	-4.397	-11.955	B10C4	-5.440	-6.161	-15.104
A9C5	-0.136	-4.136	-11.945	B10N4	-4.642	-7.213	-15.124
A9C4	-0.370	-2.776	-11.945	B10N3	-4.886	-4.940	-15.114
A9C6	-1.249	-5.031	-11.936	B10C2	-5.667	-3.827	-15.095
A9O6	-1.245	-6.261	-11.934	B10O2	-5.192	-2.681	-15.103
A9N1	-2.466	-4.339	-11.928				
A9O2	-2.602	-2.966	-11.929				
A9N2	-3.859	-2.500	-11.920				
A9N3	-1.570	-2.127	-11.937				
A10N9	1.923	-4.834	-15.373				
A10C8	2.053	-6.207	-15.357				
A10N7	0.914	-6.851	-15.309				
A10C5	-0.038	-5.837	-15.290				
A10C4	0.570	-4.600	-15.329				
A10C6	-1.463	-5.909	-15.240				
A10O6	-2.182	-6.907	-15.201				
A10N1	-2.043	-4.634	-15.239				
A10O2	-1.348	-3.443	-15.280				
A10N2	-2.092	-2.328	-15.270				
A10N3	-0.020	-3.370	-15.327				

For this set of backbone and GC base coordinates the two-fold axis, L_2, lies along the y axis and R_2 is defined by the line joining the points (-1.000, -2.311, -16.998) and (1.000, 2.311, -16.998).

2.5 AXIAL PATTERSONS

As described in §2.3, the axial Patterson is a convenient means of comparing the SBS and DH models with the fibre diffraction data. Graeme McKinnon (1980) computed axial Pattersons of the SBS and DH models using accurate representations of the electron density of all the non-hydrogen atoms in the molecule. The results are similar to those in Fig. 2.13 and still favour the SBS model because of the better agreement with the data at z = 6.8 Å (Bates *et al.*, 1980). The stereochemical refinement of the SBS model allows a more accurate comparison with the X-ray data to be made and so axial Pattersons were computed (using McKinnon's (1980) techniques) for the refined structure given in table 2.8. The axial Patterson should be computed using a model which is periodic, so the long range twist of the SBS structure must be removed before $P_{ax}(z)$ is calculated. The effect of the long range twist on the axial Patterson is removed in two different ways. In the first approach, two different axial Pattersons are computed by treating two different ten base pair sections of the SBS structure as "repeat units". One axial Patterson is computed using nucleotide pairs 1 to 10 as the repeat unit (see Fig. 2.15(a)) and the other using nucleotide pairs 6 to 5 as the repeat unit (see Fig. 2.15(b)). These two axial Pattersons are then averaged to give an axial Patterson representation of an "averaged SBS structure". In the second approach each nucleotide pair is rotated by 4.68° relative to its neighbour to produce a structure with no long range twist, allowing direct computation of the axial Patterson for an "untwisted SBS structure".

These two SBS axial Pattersons are compared with those calculated from the DH (with AT base pairs) and the fibre diffraction data (taken from Bates *et al.*, 1980) in Fig. 2.16. Inspection of Fig. 2.16 shows that there is little difference in the forms of the two curves computed from the SBS model except at z = 5 Å. The curve computed from the data has been averaged over three curves computed from three independent sets of diffraction data taken from Feughelman *et al.* (1955), Bram (1977) and Zimmerman and Pheiffer (1979). Inspection of Fig. 2.16 shows that the SBS model still fits the diffraction data better than the DH because of the better agreement at z = 6.8 Å. The small peak for the averaged SBS model at z = 5 Å is due to the phosphate groups B8 and B9 in the bend regions. Small lateral displacements of these atoms removes this peak as shown by the axial Patterson for the untwisted structure. The reduced amplitude of the SBS axial Pattersons at z = 17 Å is due to the zig-zag

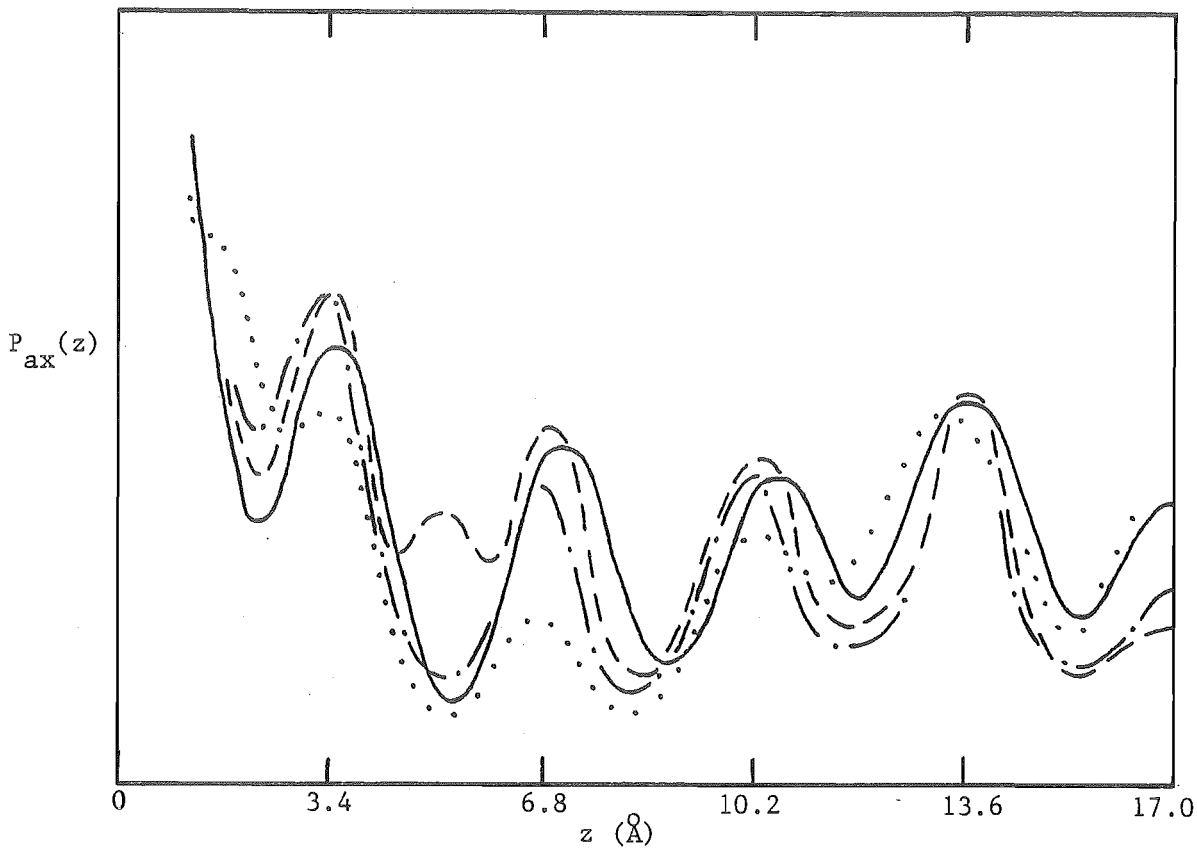


Fig. 2.16 Axial Pattersons calculated from the diffraction data (——), the DH model (••••) and the "averaged" (----) and "untwisted" (—•—•) stereochemically refined SBS models (see text). The solid and dotted curves are taken from Bates *et al.* (1980).

arrangement of the base pairs (see Fig. 2.15) so that those separated by $c/2 = 17$ Å are offset from each other. This indicates that the displacements of some of the base pairs from the molecular axis may have been overestimated. The SBS structure is probably sufficiently flexible to allow appropriate small adjustments to be made without affecting the stereochemical viability of the model.

2.6 DISCUSSION

Although the question of DNA structure is considered closed by the majority of the biochemical community, there is still no direct evidence which allows one to say with certainty that long stretches of DNA adopt the DH structure. Small, but significant, interest is still being shown in the SBS structure. Recently, different forms of DNA, such as left-handed

or "Z-DNA", have been observed by Wang *et al.* (1979) and Arnott *et al.* (1980). The two main kinds of experimental evidence used to elucidate DNA structure are the X-ray data and results from experiments with covalently closed circular (CCC) DNA.

The difficulty with the X-ray data is that it is observed at such low resolution, that the Fourier synthesis techniques described in §1.7.3 cannot be used. Hence in determining the structure from the X-ray data, many *a priori* constraints are built in. The structure refinements carried out by Arnott and Hukins (1973) using the linked atom least squares (LALS) technique have, in effect, assumed a DH structure from the outset. The R-factor they obtain (0.31) need not be considered particularly low as lower R-factors have been obtained for structures which have subsequently been found to be grossly incorrect (Donohue, 1969). The SBS and DH models are structurally very similar (although they are topologically very different because of the different association of the two strands) so it is not surprising that it is difficult to distinguish between them on the basis of the X-ray data. In fact the only major structural differences are due to the presence of left-handed regions and bend regions, the first of which would be difficult to detect from low quality X-ray data. This only leaves the bend regions which is precisely what would produce the axial Patterson peak at 6.8 Å. If this peak is not an artifact (remember that three independent sets of data give similar peaks) it clearly supports the SBS model and is difficult to explain in terms of the DH. Wing *et al.* (1980) have recently reported a high resolution structural analysis (using Fourier synthesis and isomorphous replacement techniques) of a twelve base pair unit of DNA. This fragment does have a right-handed DH structure of the sort refined by Arnott and Hukins (1973). However this does not necessarily mean that much longer lengths of DNA adopt the same structure.

Experimental studies with CCC DNA (Crick *et al.*, 1979) have also been used as evidence in support of the DH. In these experiments, short lengths of DNA are inserted in closed circles of DNA and the resulting topological (degree of intertwining) changes observed. The results can be nicely explained in terms of the DH. However, when working with such short lengths of DNA it seems possible that the structure (be it SBS or DH) could be distorted, making interpretation of the results difficult. The main support (apart from the axial Patterson) for the SBS model rests on topological arguments. DNA denatures and renatures very rapidly (Bates *et al.*, 1977) and this is difficult to understand if the strands are intertwined hundreds

of thousands of times. The separation of the two strands has been observed by electron microscopy and they appear to pare apart without any topological difficulties. Interpretation in terms of the SBS model is much simpler than for the DH. Cutting and splicing enzymes are known to be present (probably for DNA repair) but whether these are involved in replication is unknown.

While the case for the DH looks very good, there are enough difficulties with it to justify a full investigation of the SBS model.

3. SOME ASPECTS OF EXACT MACROSCOPIC INVERSE SCATTERING

3.1 INTRODUCTION

In this chapter some aspects of the inverse scattering problem for macroscopic wave motion - i.e. applied to the HE rather than the SE - are examined. The emphasis is on techniques which, in principle, provide "exact" solutions as opposed to those (such as the ones described in §1.5 and §1.6) which are based on some sort of physical approximation such as weak scattering. Significant progress with practical algorithms for exact solutions (for penetrable scatterers) has only been made for one-dimensional problems so the emphasis in this chapter is on these. The implications of extending these techniques to more than one space dimension are discussed. As described in chapter 1, both the forward and inverse scattering problems can be formulated in either the time domain (1.20) or the frequency domain (1.22). Approaches to the inverse problem in these two domains are quite different, and one aim of this chapter is to demonstrate the limitations and advantages of formulations in these two different domains.

As described in §1.9, the GL method provides an exact solution to the inverse problem for the one-dimensional SE. The GL method is formulated in the time domain and is based on the application of causality arguments. Causality can be used in a similar manner to formulate a time domain inversion algorithm for the one-dimensional HE. Because the HE is non-dispersive, this inversion technique is much simpler than the GL method and is probably more stable numerically. This algorithm is described in §3.2 in terms of EM scattering of a plane wave by a stratified medium. The implications of applying this technique to the probing of dispersive media is also addressed. Pre-processing of measurements of the scattered field is important in making the application of the inversion algorithm as stable as possible. This data pre-processing is described in §3.3. In §3.4 examples of applications to both simulated and measured data for stratified media and transmission lines are presented. An interesting extension of these ideas can be made for applications to branched duct or transmission line networks. This is also related to the situation where a duct or transmission line contains localised losses. These extensions to the algorithm, together with examples of applications to both simulated and measured data, are presented in §3.5. In §3.6 a first order solution to the inverse eigenvalue problem

introduced in §1.10 is examined. The examples presented in §3.6 highlight the advantages and limitations of approximate frequency domain methods compared to exact time domain methods. In §3.7 the methods used in this chapter are compared with each other and with some of the techniques described in chapter 1.

The new results presented in §3.2, §3.4 and §3.5 have been (Bates and Millane, 1981) and are being (Millane and Bates, 1981) published.

3.2 TIME DOMAIN MACROSCOPIC INVERSE SCATTERING

The GL method for exactly solving the inverse scattering problem for the SE is based on the causal nature of the field. It might be thought that a similar inversion procedure could be applied to the HE directly (since the field is still causal) without resorting to the protracted procedure described in §1.9.2 to make use of the GL method. That this is indeed the case seems to have been recognised only fairly recently (Lesselier, 1978; Tijhuis, 1981; Bates and Millane, 1981). These techniques have sometimes been referred to as "space-time discretization" (Lesselier, 1978; Tijhuis, 1981). In fact, application to the HE is both conceptually and procedurally much simpler than the GL method because the HE is non-dispersive. However, because the wavefront travels at a speed which is unknown *a priori* (since the refractive index is initially unknown), the solution has not been expressed in closed mathematical form. This probably accounts for the lack of attention to it in the literature as compared to the GL method. Similar techniques have been used in the determination of conductivity profiles (Bolomey *et al.*, 1981) and the shapes of perfectly reflecting scatterers (Bennett, 1981), in vocal tract analysis (Stansfield and Bogner, 1973) and in reflection seismology (Aki and Richards, 1980, ch. 12).

Consider the scattering of a plane EM wave, travelling in the $\hat{\underline{x}}$ direction, by an inhomogeneous lossless half-space confined to $x > 0$. The scattering medium is plane stratified so that

$$\left.\begin{aligned} &\varepsilon = \varepsilon_o \quad \text{and} \quad \mu = \mu_o \quad \text{for} \quad x \leq 0 \\ &\text{and} \\ &\varepsilon = \varepsilon(x) \quad \text{and} \quad \mu = \mu(x) \quad \text{for} \quad x > 0 \,. \end{aligned}\right\} \tag{3.1}$$

Let $\Psi = \Psi(x,t)$ be a single component of the electric field so that (as outlined in §1.2) Ψ satisfies the HE (1.20). The free space velocity c,

refractive index $\nu = \nu(x)$ and impedance $\zeta = \zeta(x)$ are defined by (1.19),

$$\nu = (\varepsilon\mu/\varepsilon_o\mu_o)^{\frac{1}{2}} \tag{3.2}$$

and

$$\zeta = (\mu/\varepsilon)^{\frac{1}{2}} \tag{3.3}$$

respectively. In order to formulate the inversion procedure, it is convenient to approximate ν and ζ (and hence ε and μ) by the piecewise constant distributions $\tilde{\nu} = \tilde{\nu}(x)$ and $\tilde{\zeta} = \tilde{\zeta}(x)$ where

$$\tilde{\nu} = \nu_m \quad \text{and} \quad \tilde{\zeta} = \zeta_m \quad \text{for} \quad x_{m-1} < x < x_m \ , \tag{3.4}$$

the ν_m and ζ_m are constants and $x_o = 0$. The x_m are chosen such that the electrical length of each "piece" is equal to $c\tau/2$, so that

$$(x_m - x_{m-1})\nu_m = c\tau/2 \tag{3.5}$$

where τ is a constant.

Consider the scattering of the impulsive incident field

$$\Psi_o(x,t) = \delta(x - ct) \ . \tag{3.6}$$

The resulting scattered field observed on the observation plane $x = 0$, $R(t) = \Psi_s(0,t)$, which is called here the impulse response, is given by

$$\tilde{R}(t) = \sum_{n=o}^{\infty} R_n \, \delta(t - n\tau) \tag{3.7}$$

where the R_n are constants and a tilde above a quantity denotes that it applies to the approximate profile $\tilde{\nu}$ and $\tilde{\zeta}$. Equation (3.7) is of the form of a Bremmer series. Bremmer (1951) has shown that in the limit as the approximate profile approaches the actual profile (i.e. as $\tau \sim 0$), the approximate field approaches the actual field so that

$$R(t) = \lim_{\tau \sim o} \tilde{R}(t) \ . \tag{3.8}$$

Hence a reasonable estimate of each R_n is

$$R_n = \int_{n\tau}^{(n+1)\tau} R(t)dt \ . \tag{3.9}$$

Each R_n is equal to the sum of the amplitudes, at $x = 0$ and $t = n\tau$, of an impulse which undergoes a single reflection at $x = x_n$ and those which undergo multiple reflection and emerge at the observation plane at time $t = n\tau$. The causal nature of the scattered field ensures that the multiply reflected

impulses contributing to R_n can only have penetrated the scattering region a maximum distance of $x = x_{n-1}$. Hence R_n can be written in the form

$$R_n = \Gamma_n \prod_{m=o}^{n-1} (1 - \Gamma_m^2) + g_n(\Gamma_o, \Gamma_1, \ldots, \Gamma_{n-1}) \tag{3.10}$$

where Γ_m is the reflection coefficient at $x = x_m$ for a wave travelling to the right (increasing x) and g_n is equal to the sum of the amplitudes of the multiply reflected impulses which arrive at $x = 0$ at time $t = n\tau$. It follows from (3.10) that, if Γ_m is known for $m \leq n-1$, g_n can be computed and Γ_n determined from R_n. Since $\Gamma_o = R_o$, the reflection coefficients can be sequentially estimated from the impulse response. Requiring that Ψ and $\partial\Psi/\partial x$ be continuous at $x = x_n$ gives

$$\Gamma_n = (\zeta_{n+1} - \zeta_n)/(\zeta_{n+1} + \zeta_n) \ . \tag{3.11}$$

Since ζ_o is equal to the impedance of free space, the impedance, ζ_m, of each piece can be determined from the reflection coefficients. The functional dependence of g_n in (3.10) is an expression of causality (on which the inversion procedure depends) for the reflected field. A similar relation cannot be obtained for the transmitted field. This is related to the statement made in §1.9.1 that the reflection coefficient alone, but not the transmission coefficient alone, uniquely determines the scattering matrix.

The procedure outlined above provides a solution to the inverse problem since it allows the impedance of the scattering region to be reconstructed from a measurement of the impulse response. The physical basis of the technique is apparent and it allows the limitations, *a priori* information requirements and sensitivity to noise to be readily assessed. Probably the most important limitation is that only the impedance of the scattering region can be determined from the data. It is not possible to recover the individual constitutive parameters ε and μ or the refractive index. Furthermore, because of (3.5), ζ can only be determined as a function of electrical distance (and not physical distance x) into the scattering region. Additional *a priori* information is required if these ambiguities are to be resolved. For example few materials are significantly "magnetic" so that it is often reasonable to set $\mu = \mu_o$ everywhere, in which case

$$\nu = \zeta_o/\zeta \ , \tag{3.12}$$

where ζ_o is the impedance of free space, so that ν (or ε) can be determined. Since ν is then known, the electrical distance can be converted to physical distance. In the rest of this chapter it is assumed that $\mu = \mu_o$ everywhere.

The sensitivity to noise of inverse problems is illustrated by the algorithm described above. Each successive reflection coefficient computed using (3.10) makes use of the previously reconstructed reflection coefficients and hence errors tend to accumulate, making Γ_n increasingly unreliable for larger values of n. The effect of this is that errors in the reconstructed profile tend to increase with distance from the observation plane. By using (3.5) to discretize the scattering region in constant electrical distance, the procedure is much simpler than that described by Lesselier (1978) who discretizes the scattering region in constant physical distance. This increases the instability of his procedure so that he embarks on a complicated (and, it appears, computationally expensive) correction procedure during the reconstruction to correct errors in the profile.

It is instructive to consider the possibility of developing a similar inversion procedure for a dispersive system such as the SE (1.97). In the same way as described above, consider scattering by a plane stratified potential $V(x)$ which is zero for $x \leq 0$. The potential is approximated by a piecewise constant distribution V_m, each piece being of length $c\tau/2$. The reflected field for the HE can be expressed in the form (3.7) because it is non-dispersive. This means that the propagation speed in each piece and the reflection coefficient at each interface are independent of frequency. Consequently the shape of the field does not change as it undergoes transmission and reflection. Hence the value of $R(t)$, at $t = T$ say, is made up only of disturbances which have travelled an electrical distance cT in the scattering region. This is what makes possible the straightforward interpretation of the impulse response described above. Because the SE is dispersive, the propagation speed and reflection coefficients are frequency dependent. However the wavefront travels with velocity c and hence causality should allow the V_m to be recovered from the impulse response. The impulse response can be written as

$$R(t) = \sum_{n=o}^{\infty} u_n(V_1, V_2, \ldots, V_{n+1}, t - n\tau) + \sum_{n=2}^{\infty} w_n(V_1, V_2, \ldots V_n, t-n\tau) \tag{3.13}$$

where

$$u_n(\ldots,t) = w_n(\ldots,t) = 0 \quad \text{for} \quad t < 0 \ . \tag{3.14}$$

The functions u_n and w_n represent the contribution to R(t) from single and multiple reflections respectively for which the wavefront arrives at x = 0 at time $t = n\tau$. The functional dependence of the u_n and w_n on the V_m implies that, if the u_n can be isolated from R(t), then the V_n can be reconstructed in a similar manner as described above for the HE. However, because of dispersion, identification of the individual u_n is not straightforward as they may overlap in time. The dispersion (or the effective temporal spread) of the u_n and w_n, in general, increases with n. Hence one might expect to be able to isolate u_n from R(t) for values of n less than, say, n = p. The value of p is determined by requiring that the effective temporal spread of u_n for $n < p$ is less than about τ. This would allow the determination of V_n for $n < p$. The V_n could, in principle, be determined for $n \geq p$ from the leading edge of the u_n, however this is likely to be a very noise sensitive procedure. If V(x) actually varies continuously with x then there is no obvious way of unravelling the u_n from R(t). The above discussion shows that the time domain procedure outlined for the HE is not really suitable for the SE because of dispersion. It is for this reason that one must resort to the GL algorithm in this case.

3.3 DATA PRE-PROCESSING

The inversion procedure described in §3.2 makes use of the impulse response as data. The impulse response is not a directly observable quantity (it would require an incident field of infinitely wide bandwidth) and hence has to be estimated by pre-processing the observed scattered field. The linearity of the HE implies that the scattered field observed at x = 0 takes the form

$$\Psi_s(0,t) = R(t) \otimes \Psi_o(0,t) \ . \tag{3.15}$$

Hence R(t) can be estimated by deconvolving the observed $\Psi_s(0,t)$ with $\Psi_o(0,t)$.

This pre-processing of the data is an important part of the inversion procedure presented here. In a similar time domain technique described by Lesselier (1978), samples of the profile are determined from the leading edge (in order to make use of causality) of Ψ_o. Because this leading edge is (usually) of low amplitude, the procedure is very noise sensitive and

the reconstructed profile degrades rapidly with distance from the observation plane. Tijhuis (1981) avoids using the leading edge of Ψ_o to compute the profile by formulating the problem for all of the profile so that successive samples are not determined iteratively. An initial guess is made at the whole profile which is updated in an iterative manner and hence has features of a model fitting process. However the convergence of this technique is not assured and it is only really suitable for a region of comparatively narrow width. The data pre-processing step used here avoids these difficulties and allows a simpler and more stable inversion.

Two different deconvolution methods (outlined in Appendix 1) are used for the examples presented in the following sections, depending on the type of refractive index profile. If the profile is discontinuous (i.e. piecewise constant) then R(t) consists of a train of delta functions corresponding to reflections from the discontinuities in ν. In this case the "clean" deconvolution algorithm (Schwarz, 1978) is used to deconvolve R(t) from $\Psi_s(0,t)$. Because "clean" is a subtractive technique (Bates *et al.*, 1982a, 1982c), it is suitable for deconvolving data of this type and produces a discrete impulse response. For smooth profiles, R(t) is a continuous function of time and "Wiener filtering" (McDonnell *et al.*, 1976 and Bates *et al.*, 1982b) is a suitable deconvolution procedure.

3.4 EXAMPLES OF TIME DOMAIN INVERSE SCATTERING

In this section, results of applying the technique described in §3.2 to simulated and measured data are presented. For the simulated data (generated in the computer), the incident field is a Gaussian pulse given by

$$\Psi_o(0,t) = \exp(-t^2/t_o^2) \tag{3.16}$$

where t_o is a constant and the "length", L, of the pulse is defined by

$$L = 2ct_o \ . \tag{3.17}$$

The length is approximately equal to the distance between half power points when the pulse propagates in free space. The scattered field is computed using (3.11), (3.10), (3.7) and (3.15) followed by the addition of "noise". The noise is normally distributed with an rms signal to noise ratio of $1/\gamma$. The scattered field is then deconvolved (either cleaned or Wiener filtered

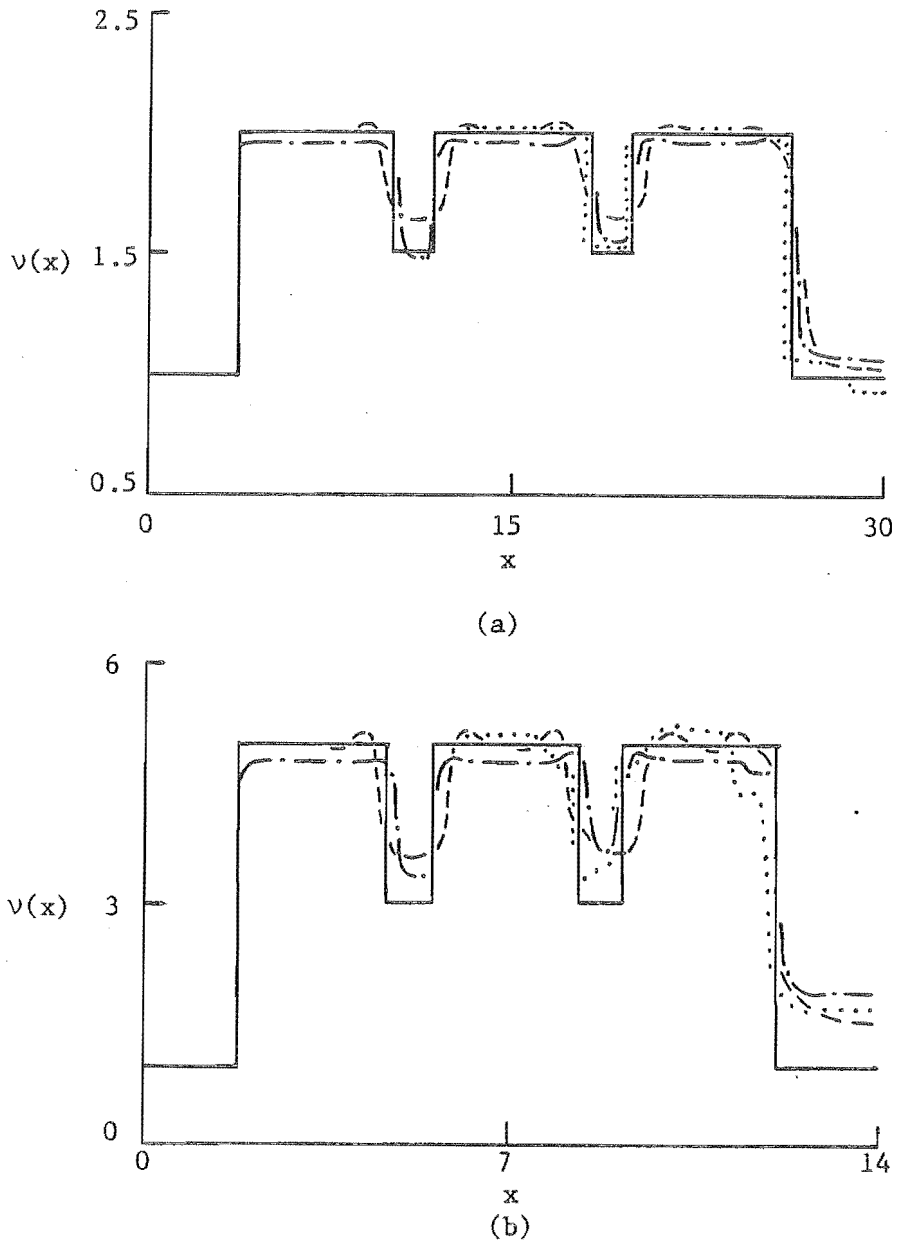


Fig. 3.1 Reconstructions of a discontinuous $\nu(x)$ from simulated data. $\gamma = 0$ and $L = 3.5$ (——) is indistinguishable from the actual profile. SSA, $\gamma = 0$ and $L = 3.5$ (····). $\gamma = 0$ and $L = 14$ (– – –). $\gamma = 0.2$ and $L = 3.5$ (–·–·).

as appropriate, see §3.3 and Appendix 1) to estimate the impulse response. The numbers R_n are estimated using (3.9) for 100 values of n. These are used in the inversion procedure to reconstruct $\nu(x)$.

The results of these simulations for discontinuous and smooth profiles are shown in Figs 3.1 and 3.2 respectively. The profiles in Figs 3.1(b) and 3.2(b) have a larger variation in ν than those in Figs 3.1(a) and 3.2(a).

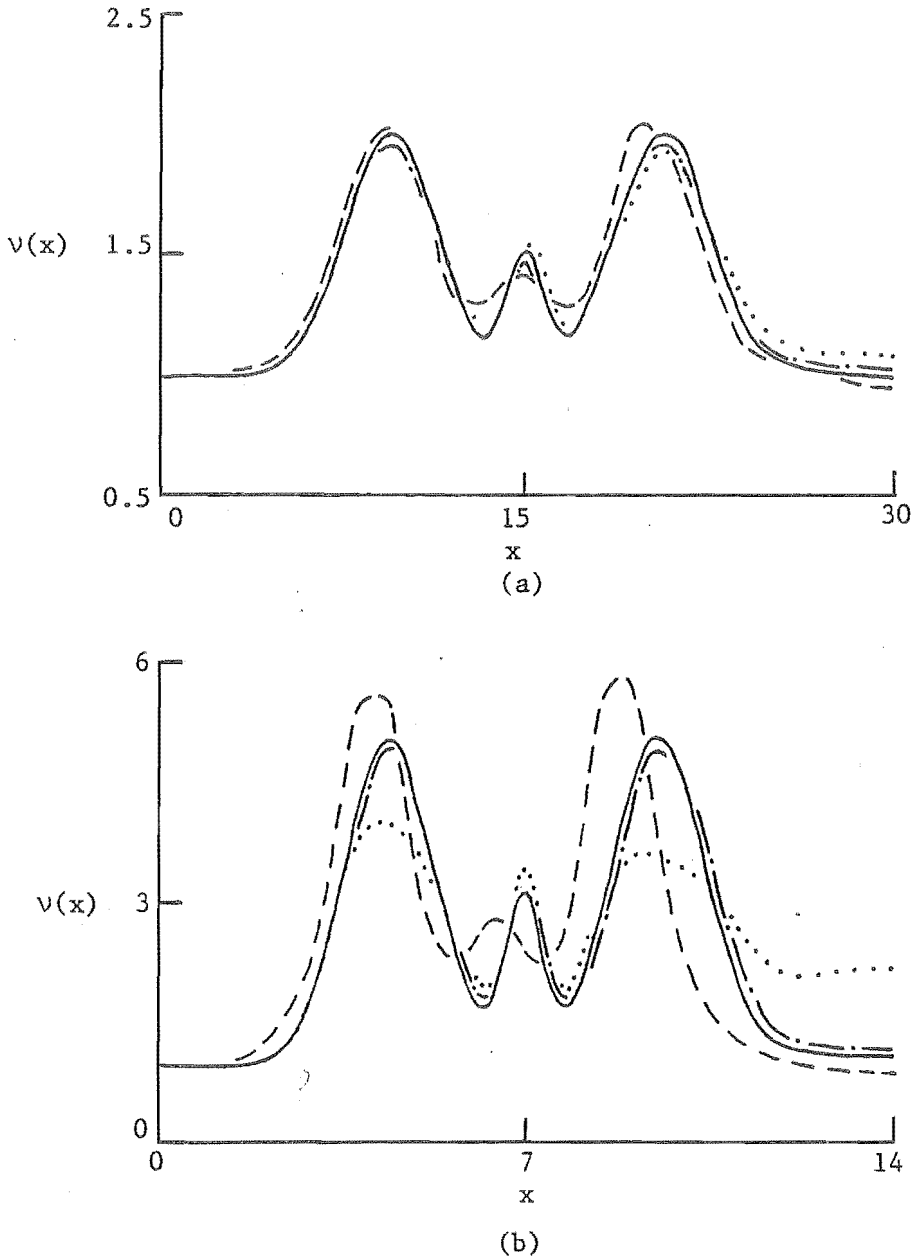


Fig. 3.2 Reconstructions of a smooth $\nu(x)$ consisting of three Gaussian inhomogeneities. The different curves are defined in the caption to Fig. 3.1.

The actual profiles (solid lines in Figs 3.1 and 3.2) are accurately reconstructed if there is no measurement noise and if the length of the incident field is less than twice the electrical length of all detail in the profile.

If the profiles were reconstructed using elementary echo-location principles then the variable propagation speed, attenuation of the incident field as it penetrates the scattering region and multiple scattering would be neglected. Such a reconstruction would be very erroneous because of the

large variations in refractive index shown in Figs 3.1 and 3.2 A more realistic approximate reconstruction is obtained by taking account of the variable propagation speed and attenuation of the incident field but neglecting multiple scattering. This is here called the single scattering approximation (SSA). The dotted curves in Figs 3.1 and 3.2 are the SSA reconstructions. Note that the deterioration in the reconstruction increases with distance from the observation plane because the number of multiple reflections increases as the field penetrates the scattering region. Inspection of Figs 3.1 and 3.2 shows that reconstructions based on the SSA are more accurate if the scatterer has a smaller variation in ν. This is to be expected, as the magnitudes of the multiple reflections are smaller. Hence, for scatterers whose refractive index variations are large, it is necessary to consider multiple scattering if accurate reconstructions are required.

The other two curves in Figs 3.1 and 3.2 illustrate the effects of the length, L, of the incident field and measurement noise γ. The electrical length of detail which is accurately resolved in the reconstructed profile is limited by L as shown in Figs 3.1 and 3.2 The dotted line in Fig. 3.2 (b) shows additional errors which are due to the large value of L. Both of these effects are due to the imperfect nature of the deconvolution process. Inspection of Figs 3.1 and 3.2 shows that contamination of the scattered field with noise produces errors in the reconstruction which increase with distance from the observation plane as described in §3.2. A comparison of these two figures suggests that the reconstructions of discontinuous profiles are more sensitive to noise than are smooth profiles. This is because the noise causes the deconvolution procedure to misposition the peaks in R(t) so that the positions of the steps in ν are slightly in error. Hence predicted multiple reflections do not accurately cancel those in the data, causing errors in the reconstructed profile.

In order to test the inversion algorithm on real data, measurements were made using a time domain reflectometer (TDR) feeding a coaxial transmission line. The experimental procedure and apparatus are described in Appendix 2. Inspection of (1.39), (1.40) and (3.12) shows that the inversion procedure described in §3.2 can be applied to a transmission line of uniform cross-section filled with a material of constant permeability, but the permittivity of which varies arbitrarily with x. The impulse response of the line is estimated by cleaning (see Appendix 1) the

reflected waveform with the incident waveform. The impulse response is processed to recover the distribution of refractive index in the line. Reconstructions are shown in Fig. 3.3. The reconstruction is reasonably accurate but deteriorates with distance x from the input to the line. A reconstruction based on the SSA is also shown in Fig. 3.3. Even for this relatively small variation in refractive index, the effect of multiple scattering has become quite significant towards the end of the line remote from the input.

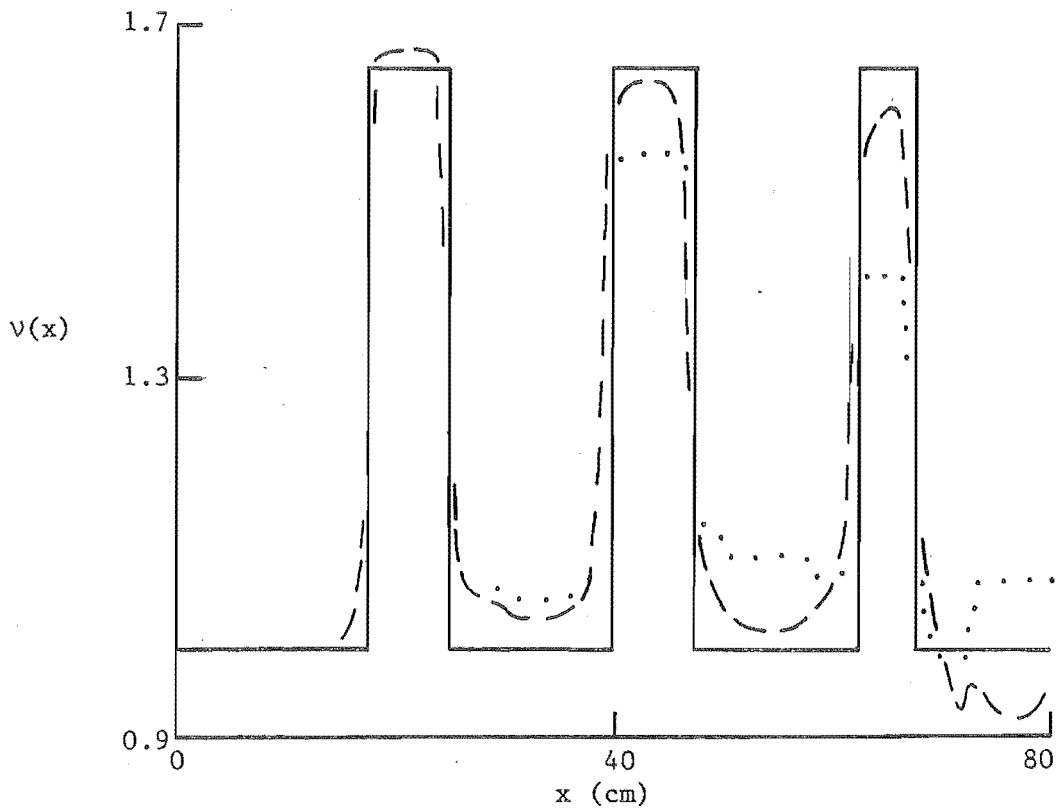


Fig. 3.3 Reconstructions of ν(x) in a coaxial transmission line from measured data obtained as described in Appendix 2. The widths of the perspex beads are 7.5 cm, 7.5 cm and 5.0 cm. Actual profile (——), reconstructed profile (– – –) and SSA reconstruction (• • •).

3.5 BRANCHED NETWORKS

In this section, inverse solutions to a class of problems intermediate between one-dimensional and many-dimensional problems, i.e. those involving branched duct or transmission line networks, are considered. As mentioned in §3.4, the inversion procedure is, under the appropriate conditions, applicable to ducts and transmission lines.

The shape (i.e. the cross-sectional area as a function of position) of the vocal tract is of fundamental importance in speech production, and methods of determining the shape using acoustical measurements made at the lips have been examined, for example, by Gopinath and Sondhi (1970) and Stansfield and Bogner (1973). However, in such analyses, the branch from the pharynx to the nasal cavity is not usually accounted for. This branch is important because motion of the velum during speech, changes the coupling to the nasal cavity (Hardcastle, 1976, ch. 5). The techniques outlined in this section provide a starting point for methods which could be used to characterise the effect of the branch to the nasal cavity. However, before meaningful results could be obtained, a more complete model of the branch would probably be required in order to usefully interpret the geometry of the branch in terms of the reflection coefficients.

Referring to Appendix 3, the inverse problem considered is to determine the position and impedance of branches off the main line and whether they are in parallel or series with the main line. The discussion in §3.2 and Appendix 3 makes it clear that this information can be recovered, using (A3.1) and (A3.2), from the impulse response measured at the observation point $x = 0$ if the "structure" (see Appendix 3) of the network is known. This is possible because the parameter to be reconstructed is uniquely related to an impedance which is a function of increasing electrical distance from the observation point. It is important when reconstructing branched networks to neglect responses in $R(t)$ which are below a threshold (based on the estimated noise level) to minimise the number of false branches produced in the reconstruction. Apparent energy loss into false branches rapidly degrades the reconstruction. For the same reason, the electrical distance between junctions in the network must be greater than $L/2$ so that errors introduced by the deconvolution procedure do not introduce false branches into the reconstruction. If there are a small number of possibilities for the structure, they can each be tried in turn. Some of the structures chosen which are incorrect may lead to an inconsistency during the subsequent reconstruction procedure by predicting a reflection coefficient at some point in the network which has a magnitude greater than unity. If it is known *a priori* that the network contains no closed loops then these structures can be rejected from the set of possibilities considered. In order to detect incorrect structures in this fashion it is essential to take the multiple reflections into account to make the estimation of the reflection coefficients as accurate as possible.

If the observed data allows more than one consistent solution then the "simplest" network is "probably" the correct solution. The simplest network is defined as being the one with the smallest number of branches or impedance changes. This is because, if the predicted branching pattern is incorrect, single reflections are interpreted as originating from wrong positions in the network. Hence some of the multiple reflections predicted will not cancel those in the data and will be interpreted as originating from additional branches or impedance changes. For this reason it seems sensible in many practical situations to choose the reconstruction which is the simplest network.

As intimated in Appendix 3, the magnitude and location of localised losses in the network can also be determined in the same manner as described above. Of course a localised loss cannot be distinguished from an equivalent infinite branch unless some *a priori* information is available that only one or the other is present. Inspection of (A3.3) indicates that it is possible to distinguish between parallel and series losses. Experimentally determined vocal tract area functions show a sharp discontinuity at the epiglottis (Gopinath and Sondhi, 1970) which would lead to viscous energy loss at this point. It may be possible to take this loss into account using the techniques described here.

Results are presented of applying the ideas described above to both simulated and measured data. In order to obtain sensible results, a particularly simple branching pattern is considered. The main line is of constant known impedance ζ_o with single branches terminated in matched loads. The inverse problem is to recover the positions and impedances of the branches, and whether they are in series or parallel with the main line, from an estimate of the impulse response measured at one end of the line.

Numerical simulations were performed to illustrate the effects of multiple scattering and the possibility of distinguishing between different branching patterns. The two different branching patterns considered are that described above and an unbranched line simply consisting of impedance changes. In these examples both the noise level and incident signal length are assumed to be very small. Fig. 3.4(a) shows a branched line which is reconstructed when multiple scattering is taken into account. Fig. 3.4(b) shows a reconstruction based on the SSA. The neglect of multiple scattering causes errors in the estimate of the branch impedances and introduces false branches into the line. Fig. 3.4(c) shows the effect of processing the data on the basis of an unbranched line with a variable

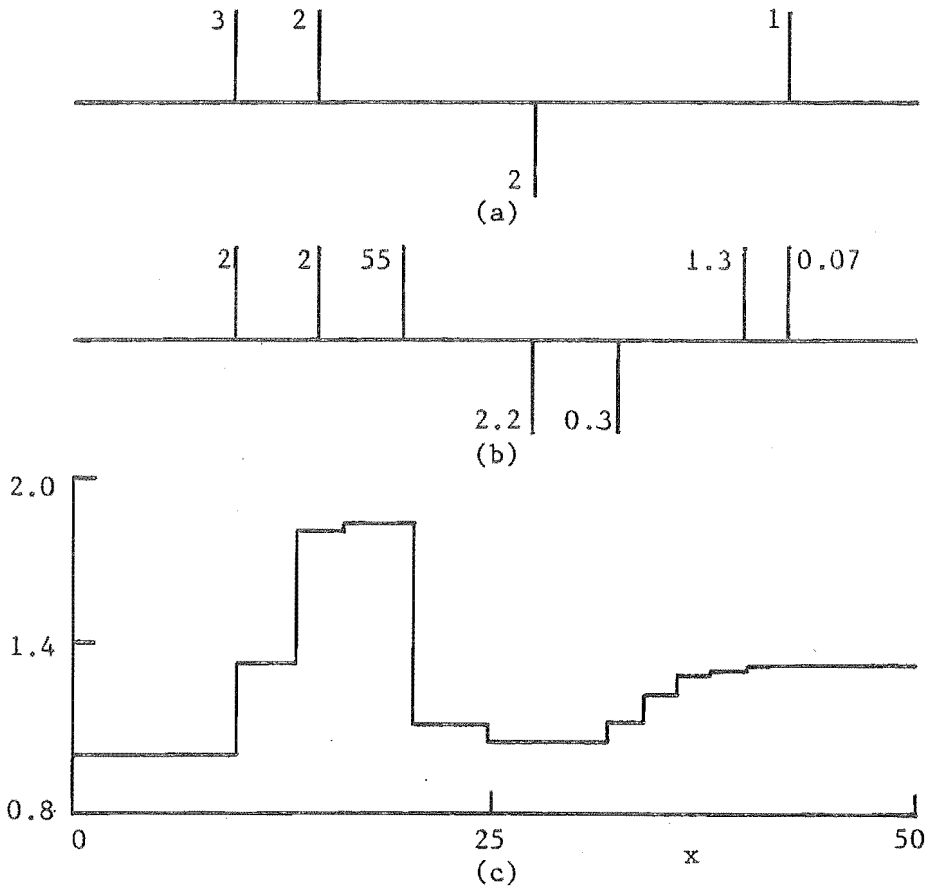


Fig. 3.4 Reconstructions of a branched network from simulated data (with $\gamma = L = 0$). Vertical lines above and below the horizontal line indicate parallel and series branches respectively. The numbers next to each branch denote ζ/ζ_0 where ζ and ζ_0 are the impedances of the branches and the main line respectively.
(a) Actual branched network which is reconstructed when multiple scattering is taken into account.
(b) SSA reconstruction.
(c) Profile of $\nu(x)$ obtained when data are interpreted as pertaining to an unbranched line.

refractive index. The simplicity of Fig. 3.4(a) over Fig. 3.4(c) suggests that the line does in fact contain branches. Fig. 3.5 shows the result of processing the data from Fig. 3.1(a), on the assumption that it originates from a branched network. The arrow in Fig. 3.5 shows the point at which the reconstruction procedure indicates that the data is inconsistent as described above. This shows that the data does not originate from a branched network of this type.

The inversion procedure for branched networks was applied to measurements made using the TDR (see Appendix 2) feeding branched transmission line networks. The networks were constructed from coaxial sections and T-junctions. The impulse response was estimated by cleaning (see Appendix 1) the measured reflected waveform. Fig. 3.6(b) shows a

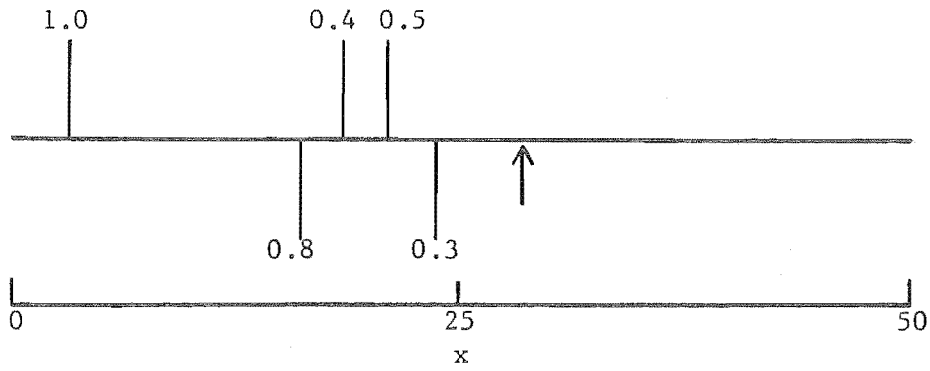


Fig. 3.5 Reconstruction of a branched network using the data characterised by the full line in Fig. 3.1(a). The arrow indicates the point where, during the inversion procedure, the data become inconsistent. The numbers next to each branch denote ζ/ζ_o (see caption to Fig. 3.4).

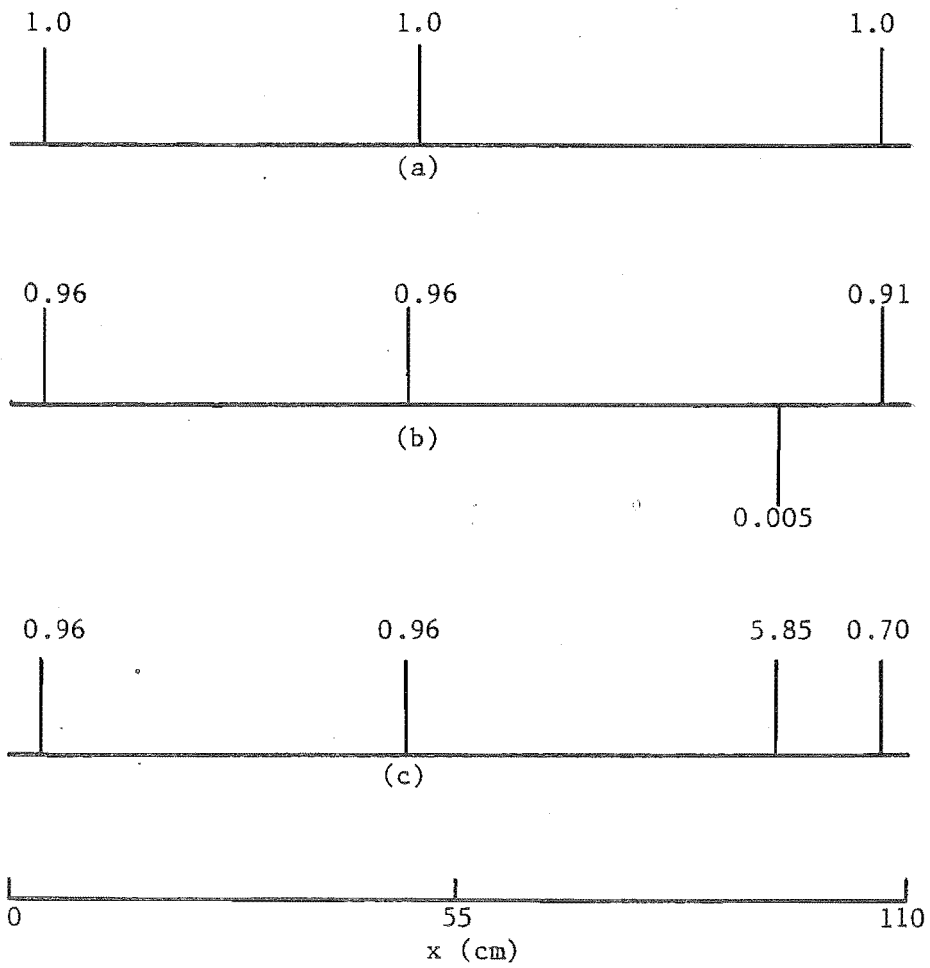


Fig. 3.6 Reconstructions of a branched network from measured data obtained as described in Appendix 2. The numbers next to each branch denote ζ/ζ_o (see caption to Fig. 3.4).

(a) Actual branched line.
(b) Reconstruction.
(c) SSA reconstruction.

reconstruction of the branched line shown in Fig. 3.6(a). Fig. 3.6(c) shows the decrease in accuracy of the reconstruction if the data is processed using the SSA.

3.6 SOLUTIONS TO THE INVERSE EIGENVALUE PROBLEM

In this section, a first order solution to the inverse eigenvalue problem introduced in §1.10 is examined. The procedure follows closely that of Schroeder (1967) and Mermelstein (1967) used to estimate vocal tract area functions. Since this is an approximate frequency domain method, the results presented in this section, when compared with those in §3.4, provide a means of comparing, firstly, approximate and exact inverse scattering techniques and, secondly, frequency domain and time domain methods. It is shown in §1.10 that the refractive index profile ν is uniquely determined by the two spectra $\{k_n\}$ and $\{\ell_n\}$. Perhaps the most obvious way to estimate ν from these two spectra, is to compute the reflection coefficient using (1.139), Fourier transform it to obtain the impulse response and then apply the inversion procedure described in §3.2. However a large number of the eigenvalues are needed in order to compute the impulse response with reasonable accuracy. As described in §3.2, any errors in the impulse response cause a rapid degradation in the reconstruction with increasing distance from the point where it is measured ($x = 0$ in this case). In many practical situations only a small number of the resonant frequencies may be available for measurement. It is important to determine the limitations imposed by a small data set and to devise algorithms which are suitable in such situations. Hence a more direct method which does not make use of the impulse response (and is hopefully, therefore, more stable) is desirable. Since in many practical situations only one (rather than two) spectra are available, it is of interest to determine what, if any, useful information is provided by one spectrum. Backus and Gilbert (1968) have used their method (see §1.11) to estimate profiles from natural frequency data. However, although their method is extremely useful in practice, as with many model fitting procedures, it does not provide much in the way of physical insight into the problem.

Exact methods of proceeding directly from the eigenvalues to the profile (for example the method of Krein (1952) and that described in Appendix 4) are not amenable to numerical computation, so it is instructive

to consider the case where the actual profile, $\nu^2 + \delta\nu^2$, is a small perturbation $\delta\nu^2$ to a known profile ν^2. Multiplying (1.130) by ψ_n and integrating over (0,a), integrating by parts and making use of (1.131) and (1.132) gives

$$k_n^2 \langle \nu^2 \psi_n^2 \rangle = \langle \psi_n'^2 \rangle \tag{3.18}$$

where the notation $\langle f \rangle$ is defined by

$$\langle f \rangle = \int_0^a f(x)\, dx \tag{3.19}$$

and a prime denotes the derivative with respect to x. Considering small perturbations δk_n^2 and $\delta\psi_n$ to k_n^2 and ψ_n, respectively, evaluating (1.130) to first order and making use of (1.131), (1.132) and (3.18) gives

$$\delta k_n^2 \langle \nu^2 \psi_n^2 \rangle = -k_n^2 \langle \delta\nu^2 \psi_n^2 \rangle \ . \tag{3.20}$$

Similar reasoning shows that

$$\delta \ell_n^2 \langle \nu^2 \phi_n^2 \rangle = -\ell_n^2 \langle \delta\nu^2 \phi_n^2 \rangle \ . \tag{3.21}$$

Note that $\delta\psi_n$ and $\delta\phi_n$ do not appear in (3.20) and (3.21). Equations (3.20) and (3.21) are integral equations for $\delta\nu^2$ in terms of the eigenvalues. Note that these integral equations are of the first kind Fredholm type (Porgorzelski, 1966) and hence are inherently unstable with respect to noise or incompleteness in the data.

A stable way of making use of (3.20) and (3.21) is as follows. It can be shown (Barcilon, 1979) that the functions ψ_n^2 and ϕ_n^2 together form a complete orthogonal set on the interval (0,a), which suggests that it might be useful to expand $\delta\nu^2$ in terms of these functions. Consider the case where the unperturbed profile is the constant value $\nu^2 = A_o$ so that the eigenfunctions are given by

$$\left.\begin{aligned} &\psi_n = \sin(n\pi x/a) \\ &\text{and} \\ &\phi_n = \cos\left((n-\tfrac{1}{2})\pi x/a\right) \ . \end{aligned}\right\} \tag{3.22}$$

The appropriate expansion for $\delta\nu^2$ is then the Fourier cosine series

$$\delta\nu^2 = \sum_{n=1}^{\infty} \{A_n \cos(2\pi n x/a) + B_n \cos((2n-1)\pi x/a)\} \ . \tag{3.23}$$

Substituting (3.22) and (3.23) into (3.20) and (3.21) and using the orthogonality of cos and sin gives

$$\left.\begin{aligned} &\delta k_n^2/k_n^2 = A_n/2A_o \\ \text{and} \\ &\delta \ell_n^2/\ell_n^2 = -B_n/2A_o \ . \end{aligned}\right\} \quad (3.24)$$

In order to make use of (3.24) the value of $A_o = \langle \nu^2 \rangle$ and the unperturbed eigenvalues k_n^2 and ℓ_n^2 must be estimated. Let $\tilde{k}_n^2$ and $\tilde{\ell}_n^2$ be the measured eigenvalues which are measured for $1 \leq n \leq N$. Making use of (3.23) and remembering that for small perturbations $A_n << A_o$ and $B_n << A_o$, then to first order

$$A_o = \langle \nu \rangle^2 \ . \quad (3.25)$$

Hence, using (1.134), (1.135), (1.137) and (3.25), allows A_o to be estimated by

$$\left.\begin{aligned} &A_o \simeq (N\pi/\tilde{k}_N a)^2 \\ \text{or} \\ &A_o \simeq \left((N-\tfrac{1}{2})\pi/\tilde{\ell}_N a\right)^2 \ . \end{aligned}\right\} \quad (3.26)$$

Making use of (1.134) and (1.135) allows the unperturbed eigenvalues to be estimated by

$$\left.\begin{aligned} &k_n \simeq n\tilde{k}_N/N \\ \text{and} \\ &\ell_n \simeq (n-\tfrac{1}{2})\tilde{\ell}_N/(N-\tfrac{1}{2}) \ . \end{aligned}\right\} \quad (3.27)$$

Equations (3.23), (3.24), (3.26) and (3.27) allow the profile to be estimated from the measured eigenvalues. Equation (3.24) confirms (at least to first order) the observations made in §1.10, that both sets of eigenvalues are required to specify ν and that only the set $\{\tilde{k}_n\}$ is required if it is known *a priori* that ν is symmetric about $x = a/2$, i.e. all the $B_m = 0$.

The results of computer simulations of the above procedure are shown in Figs 3.7 and 3.8. The refractive index profiles have been estimated from N of each of the eigenvalues $\tilde{k}_n$ and $\tilde{\ell}_n$. Because they have been determined using a first order procedure, the reconstructions of profiles with larger variations in ν are less accurate than those with smaller variations. However, even for the larger variations in ν, the

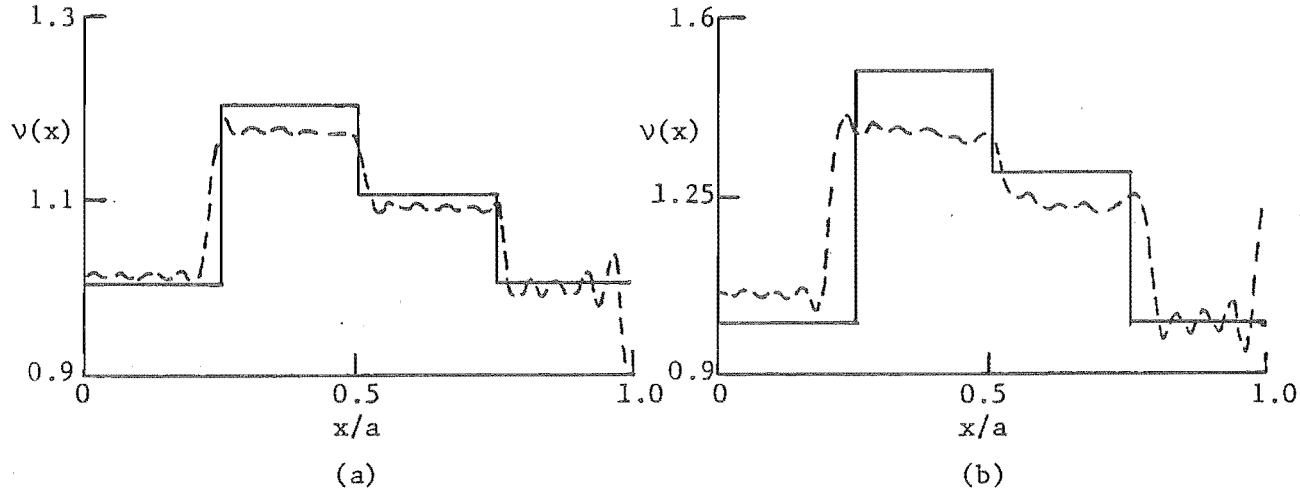


Fig. 3.7 Reconstructions (---) of discontinuous refractive index profiles (——) for N = 20.

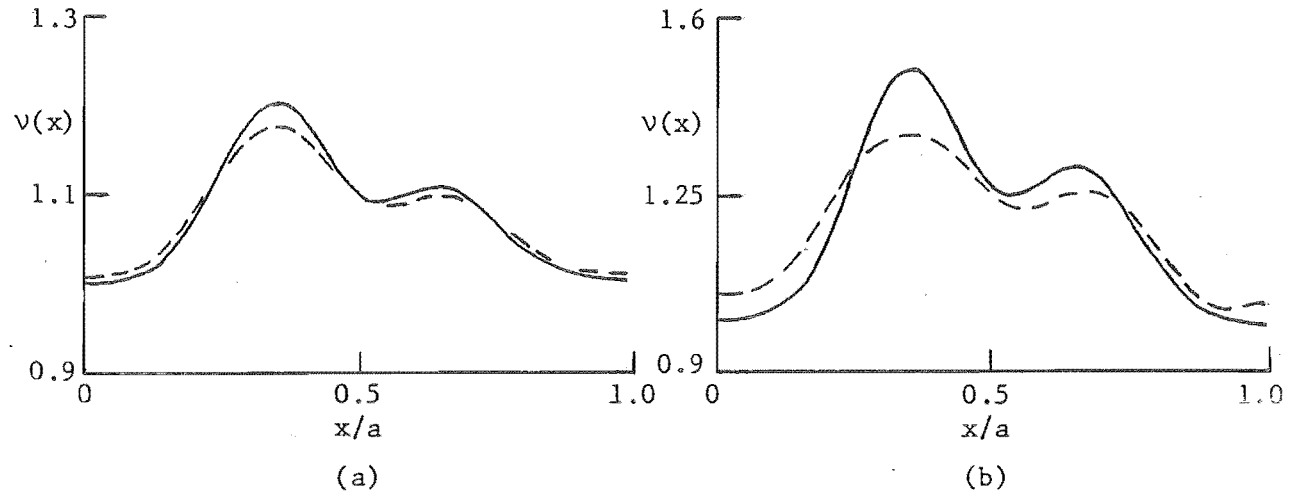


Fig. 3.8 Reconstructions (---) of smooth refractive index profiles (——) for N = 10.

reconstructions give a reasonably faithful estimate of the shapes of the profiles. Note, however, that the variation in ν in Figs 3.7 and 3.8 is much smaller than that in Figs 3.1 and 3.2 The small number of resonant frequencies used has limited the resolution of the discontinuities in the reconstructed profiles shown in Fig. 3.7. Probably the most significant feature of the reconstructions in Figs 3.7 and 3.8 compared to those in Figs 3.1 and 3.2 is that the errors in the former are spread throughout the profile whereas the errors in the latter increase away from x = 0. This feature is a result of the difference between utilising frequency domain and time domain techniques. Mermelstein (1967) has been able to improve reconstructions of vocal tract area functions by using approximate

profiles reconstructed as above as starting points in an iterative model fitting type of procedure. He appears to have obtained good results but convergence is not assured. It may be possible to extend the first order solution of the inverse eigenvalue problem described here to more dimensions by making use of appropriate eigenfunction expansions.

3.7 DISCUSSION

The time domain approach to inverse scattering described in §3.2 provides an exact solution to the one-dimensional problem for the HE. The physical basis of the method is readily apparent which allows one to assess what information can and cannot be recovered from a set of measurements and what *a priori* information is required in a particular application. The procedure is potentially noise sensitive but provides accurate results if one does not attempt to probe the medium too far from the measurement point. By deconvolving the scattered field and basing the inversion on the impulse response, the procedure is more stable than if the scattered field is used directly. By sampling the profile in constant electrical, rather than physical, distance the procedure is computationally more efficient and more stable than that reported by Lesselier (1978). The reconstructions shown in §3.3, calculated from simulated data, show that useful reconstructions can be obtained even in the presence of a reasonably high level of measurement noise. The experimental results show that the technique is sufficiently stable to provide useful results from measured data.

The time domain method is closely related to the GL method for the SE. It is simpler to apply to regions of variable refractive index than the procedure described in §1.9.2 which makes use of the Chandrasekhar transform and the GL method. However the method is not, in general, suitable for dispersive media as outlined in §3.2. The reconstructions in the examples show the importance of considering multiple scattering even if changes in impedance in the scatterer are quite small. If elementary echo-location ideas were used directly, the results would be even worse than the dotted lines in Figs 3.1 - 3.3 because of the effect of delay and attenuation of the incident field. This would be similar to using the inverse Born approximation described in §1.5 in which the incident field is assumed to propagate through the scattering region as if it were propagating through free space. If the SSA is used then

the reconstructions may still contain significant errors. This is equivalent to using only the first term of the Bremmer series for the wavefunction inside the scatterer. It is also equivalent to the first order WKB forward scattering approximation (Bremmer, 1951). However no frequency domain "inverse WKB" method has yet been formulated. It would be possible to construct a time domain inversion procedure which takes account of a finite number, N say, of terms in the Bremmer series (i.e. N = 3 includes single, double and triple reflections), the accuracy of the procedure increasing with N. This may be more sensible in some practical situations than taking N to be indefinitely large as has been done, in principle, in the examples here. Similar frequency domain inverse methods of higher order than the Born approximation have so far eluded discovery.

The first order solution to the inverse eigenvalue problem described in §3.6 illustrates a number of features common to approximate frequency domain inverse scattering techniques. Similar results would be obtained if the Born approximation were used with the total field inside the scatterer replaced by $\exp(ik\langle\nu\rangle x)$ rather than $\exp(ikx)$. The method is inherently more stable than the time domain method as successive samples in the reconstruction do not depend on previously calculated samples. Hence the errors tend to be spread throughout the profile rather than increasing with distance from the observation point as occurs when using the time domain technique. However, because the method is approximate, it only provides useful results where the variation in refractive index is small.

Attempts to extend exact macroscopic inverse scattering solutions for one-dimensional problems to more than one dimension have been largely unsuccessful. Frequency domain formulations of inverse scattering tend to be more easily extended to more than one dimension. However, useful frequency domain methods are only approximate. The difficulty with extending exact time domain techniques to more than one dimension is that the positions of the wavefronts are unknown *a priori* which makes it difficult to apply causality conditions. Bates and Millane (1981) propose an extension of the time domain method outlined in §3.2 to three dimensions by using a three-dimensional form of the Chandrasekhar transform to construct a fictitional space in which the wavefront travels with a constant speed. However scattering in this fictitional space is dispersive and so there are non-trivial difficulties, as described in §3.2, in interpreting the scattered field. The GL method has been extended to three dimensions (see

§1.9.3) using spectral techniques. As H.E. Moses intimates on page 8-33 of Colin (1972), an interpretation of the three-dimensional GL method in terms of causality would probably provide significant insight into the multi-dimensional inverse scattering problem. The practical implications of the three-dimensional GL method are as little understood as is the possibility of applying it to the three-dimensional HE.

The extension of the time domain technique to branched transmission line networks is possible because the location of the wavefront is predetermined. Localised losses may be handled by the same technique which has potential application in vocal tract analysis. It is shown in §3.5 that, in some circumstances, it may be possible to determine the correct branching pattern from a number of possibilities by requiring consistency of the data and simplicity of the network. For ease of exposition, the branched systems examined in §3.5 have a particularly simple branching pattern. However the technique is applicable to any branched or tree-type duct structure, provided that it does not contain closed loops. For a complicated branching pattern, considerable *a priori* information would be needed to allow a unique reconstruction from the impulse response measured at only one point in the network. Nevertheless, the requirements of consistency and simplicity are likely to reduce the amount of necessary *a priori* information. Also, if the impulse response was measured at more than one point in the network, it might be possible to construct a more nearly unique solution relying on less *a priori* information.

4. CARDIAC ELECTROPHYSIOLOGY

4.1 INTRODUCTION

Disorders of cardiac rhythm (known as cardiac arrhythmias) are potentially lethal, particularly in patients with some types of heart disease (Krikler and Goodwin, 1975). Recent advances in the electrophysiology of cardiac tissue (Nayler and Krikler, 1975) and the development in the 1960s of safe techniques for recording intracardiac electrical activity (Curry, 1975) have led to a greater understanding of arrhythmias. In addition, intracardiac stimulation (or "pacing") (Curry, 1975), besides being a means of initiating and terminating arrhythmias, has allowed accurate measurement of parameters of cardiac conduction, thereby enhancing the understanding of arrhythmic mechanisms. This has allowed effective treatment to be chosen for many arrhythmias in the form of surgery (Spurrell, 1975a), and the use of cardio-active drugs (Dreifus *et al.*, 1976) and pacemakers (Spurrell, 1975b). Electrophysiologic testing is comprehensively reviewed by Fisher (1981).

In chapter 5, a computer model used to assist in the analysis of cardiac arrhythmias is described. This chapter serves as an introduction to cardiac electrophysiology, cardiac arrhythmias and the measurements on which arrhythmia diagnosis is based. The physiology of the heart and electrophysiology of the cardiac conduction system are outlined in §4.2. The different types of cardiac arrhythmias are described in §4.3. In §4.4 and §4.5, the main electrophysiologic measurements, the surface and intracardiac electrocardiograms respectively, are described.

4.2 CARDIAC PHYSIOLOGY

The heart is a vital organ which acts as a pump to circulate blood throughout the body. It consists of four chambers called the left and right atria and the left and right ventricles (see Fig. 4.1). The right and left ventricles are the main pumping chambers and they circulate blood through the lungs and the rest of the body, respectively. The left and right atria pump blood into, or "prime", the left and right ventricles respectively. The four chambers pump the blood by contraction of the cardiac muscle tissue or myocardium which surrounds each chamber.

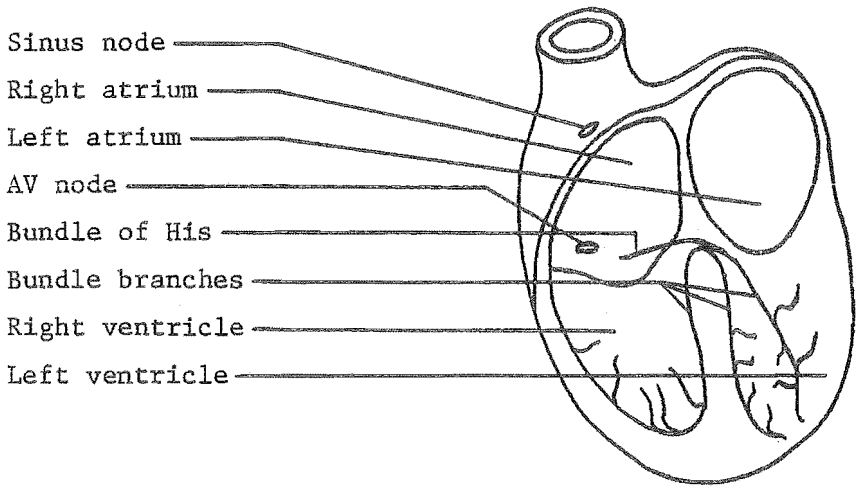


Fig. 4.1 The components of the cardiac conduction system.

The cells which make up most of the heart tissue are electrically excitable. Each cell is enclosed in a membrane and it is the flow of ions (particularly Na^+, K^+ and Ca^{2+}) across the cell membrane which is the basis of its excitory properties. The electrical state of a cell is described by the transmembrane potential or action potential (AP) which is the voltage across the cell membrane (the inside of the cell relative to the outside). The transmembrane potential of a cell in the "resting" state is approximately -90 mV so the cell is normally polarised. Cell activation is premediated by a change in the membrane permeability to certain ions. Ions then flow across the membrane and the cell depolarises to a potential of approximately zero mV (see Fig. 4.2). A cell depolarises when it receives a sufficiently strong electrical stimulus, either externally or from the depolarisation of adjacent cells.

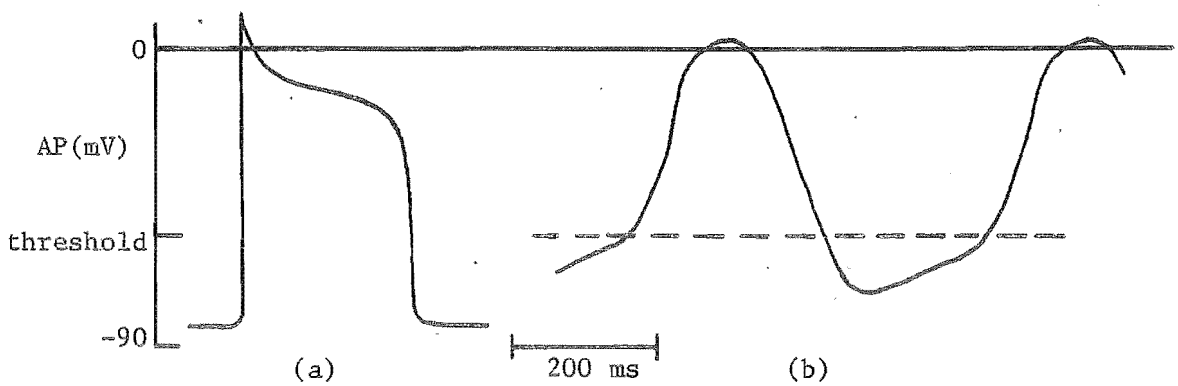


Fig. 4.2 Schematic representation of cardiac action potentials for (a) a muscle cell and (b) a pacemaker cell.

The cell remains depolarised for approximately 250 ms (during which time muscle contraction occurs) and then repolarises to the resting state. While a cell is depolarised, it is refractory, which means that it is unable to be re-activated. Some cardiac cells, called pacemaker cells, possess the property of spontaneous depolarisation. When these cells are in the resting state, the AP slowly increases until, when it reaches a threshold value, the cell spontaneously depolarises as shown in Fig. 4.2(b).

The timing of the atrial and ventricular contractions (the heart rhythm) is controlled by the propagation of action potentials (often called simply impulses) through the cardiac conduction system (CCS). The CCS consists of the sino-atrial node (SAN), atrial myocardium, atrio-ventricular node (AVN), bundle of His (HB), bundle branches and ventricular myocardium as shown in Fig. 4.1. The bundle branches divide into three divisions, of which the left anterior (LABB) and left posterior (LPBB) bundle branches terminate on the left ventricle, and the right bundle branch (RBB) which terminates on the right ventricle. The atrial and ventricular myocardia are separated by a sheet of inexcitable tissue called the AV ring. The HB provides the only electrical connection (through the AV ring) between the atrial and ventricular myocardium. The left and right atria are formed by a continuous piece of myocardium and hence depolarise together. The same applies to the left and right ventricles. The HB and bundle branches form the ventricular specialised conduction system (VSCS) and the AVN and VSCS form the atrio-ventricular conduction system (AVCS). Pacemaker cells are located in the SAN and AVCS.

The SAN is embedded in the top of the right atrium and normally provides the pacemaking impulses to the heart. These impulses propagate through the atrial myocardium (producing contraction of the atria) to the AVN which is located at the base of the right atrium. The impulses pass through the AVN, the HB and the bundle branches which spread over the ventricular myocardium. Propagation through the AVN is slow - it takes between 50 and 150 ms. This delay allows time for blood to be ejected from the atria into the ventricles before ventricular contraction occurs. The AVN has important properties which ensure the maintenance of correct cardiac rhythm. If, for some pathological reason, the atrial tissue sustains a high rate of activation, the refractory properties of the AVN protect the ventricles from the same dangerously high rate by blocking conduction of some of the impulses. If the atrial impulses fail to occur

or do not conduct to the ventricles at least about 40 times per minute, then pacemaker cells in the AVCS spontaneously depolarise to maintain a low rate of ventricular contraction. Conduction in cardiac tissue can occur in both the antegrade (from the atria towards the ventricles) and retrograde (from the ventricles towards the atria) directions.

A convenient way of illustrating conduction sequences in the CCS is by use of a ladder diagram. Fig. 4.3 shows ladder diagrams of two conduction sequences (which are described further in §4.3) through the atria (A), AVN and HB. The different sections of the CCS are represented as horizontal rows and time increases from left to right. The propagation of impulses is represented by lines through the sections, and a blocked impulse is denoted by a bar perpendicular to a line. The slope of a line gives an indication of the propagation speed.

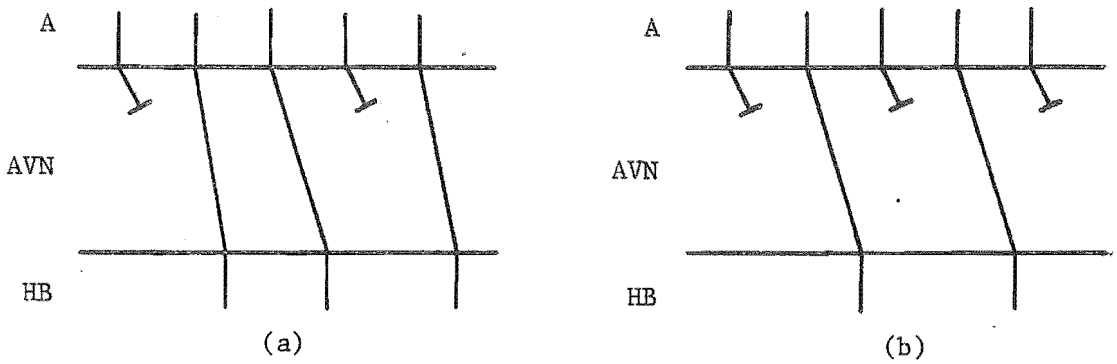


Fig. 4.3 Ladder diagrams of second degree block (see §4.3).
(a) 3:2 Wenckebach (type 1 block).
(b) 2:1 conduction (type 2 block).

4.3 CARDIAC ARRHYTHMIAS

Any defect of normal cardiac rhythm is called an arrhythmia. Arrhythmias are usually caused by disease in the CCS, electrolytic disturbances or congenital anatomical defects. Arrhythmias can be divided into those due to abnormalities of impulse production and impulse conduction.

Examples of abnormalities of impulse production are sinus arrest (no SAN activity) and sinus tachycardia (enhanced SAN activity). Excessive refractoriness of the tissue surrounding the SAN may block some of the sinus impulses from propagating to the atria. If sinus impulses do not reach the ventricles, spontaneous depolarisation usually occurs in the AVCS

and this is called a junctional rhythm. Myocardial cells (either in the atria or the ventricles) may also spontaneously depolarise to produce ectopic beats. Ectopic beats may occur either singly or in rapid succession as a tachycardia (a continuous series of ectopics from a single group of cells) or as fibrillation (a random sequence of ectopics from many sites which produces no coordinated contraction of the muscle).

Abnormalities of impulse conduction may be divided into those due to slowed conduction (usually caused by disease) and those due to the existence of an extra conducting or accessory pathway (or pathways). Slowed conduction is referred to as block and is described as occurring in three degrees (Schamroth, 1971). In first degree block, all impulses are conducted (but slowly) to their destination; in second degree block some impulses are not conducted because they encounter refractory tissue; and in third degree (or complete) block, no impulses are conducted through the site of the block. Conduction block usually occurs in the AVCS. Block in one or more of the bundle branches is referred to as bundle branch block.

First and second degree block usually occur in the AVN. Second degree block is divided into two types called type 1 and type 2 as shown in Fig. 4.3. Type 1 block (often referred to as Wenckebach - see Simpson *et al.*, 1978) is a repeating sequence, each sequence consisting of a number of conducted stimuli (each consecutive stimulus having an increased conduction time through the AVN) followed by one blocked stimulus. Type 2 block consists of a repeating sequence of only one conducted stimulus followed by a number of blocked stimuli. The conduction ratio of second degree block is defined as m:n where m and n are the numbers of applied stimuli and conducted stimuli respectively. Conduction ratios often observed during second degree block are 4:3 and 3:2 (type 1 block) and 2:1 and 3:1 (type 2 block). As the rate of stimulation increases, the AVN becomes more refractory and conduction proceeds from normal conduction, to first degree block, to second degree block. During second degree block, n/m decreases with increasing stimulation rate.

Abnormalities caused by the existence of an accessory pathway (usually between the atria and ventricles, bypassing the AVN) are called pre-excitation syndromes (Gallagher *et al.*, 1978). The accessory pathway usually conducts more rapidly than the AVN and so the ventricles are activated early - hence the term "pre-excitation". Pre-excitation syndromes may lead to tachycardias (rapid heart rates) by allowing an

impulse to circulate continuously through the atrial and ventricular tissue. An example of a pre-excitation syndrome called the "Wolff-Parkinson-White syndrome" is examined in detail in §5.5.

4.4 THE ELECTROCARDIOGRAM

The propagation of action potentials in the heart produces a time-varying electric field in the heart tissue. This causes current to flow in the tissue surrounding the heart, producing a time-varying electric field over the body surface. This time-varying body surface field is called the electrocardiogram (ECG). The ECG is measured as the differential voltage between two electrodes (or sets of electrodes) on the body surface. The ECG is the first-line diagnostic measurement used by a physician to examine cardiac arrhythmias. ECG measurements are non-invasive, safe, repeatable and relatively inexpensive. Because the ECG is so conveniently measured, an enormous amount of empirical information has been accumulated over the years relating features of the ECG to various types of arrhythmia - see, for example, Schamroth (1971). However, the ECG does not always provide enough information for the physician to determine the mechanism of an arrhythmia or the best treatment, so that the more extensive measurements described in §4.5 must be made.

The ECG of a normal person consists of three major features called the P wave, QRS wave and T wave as shown in Fig. 4.4(b). These three waves are caused by depolarisation of the atrial myocardium and depolarisation and repolarisation respectively of the ventricular myocardium. The exact shape of the ECG, of course, depends on the positions of the recording electrodes. The ECG illustrated in Fig. 4.4 (b) is the sort of signal that would be measured on lead II (electrodes positioned on the left foot and the right arm - refer to Goldman (1976), ch. 1, and Bones (1981) for a description of the different electrode positions used). The duration of the P wave averages about 100 ms which is the time taken for all of the atrial myocardium to depolarise. The interval between the onset of the P wave and the onset of the QRS wave (called the P-R interval - see Fig. 4.4(b)) is highly variable - between about 100 ms and 200 ms. This wide variation is due to the variable propagation time through the AVN which is further discussed in §4.5. The duration of the QRS wave is approximately 80 ms which is the time

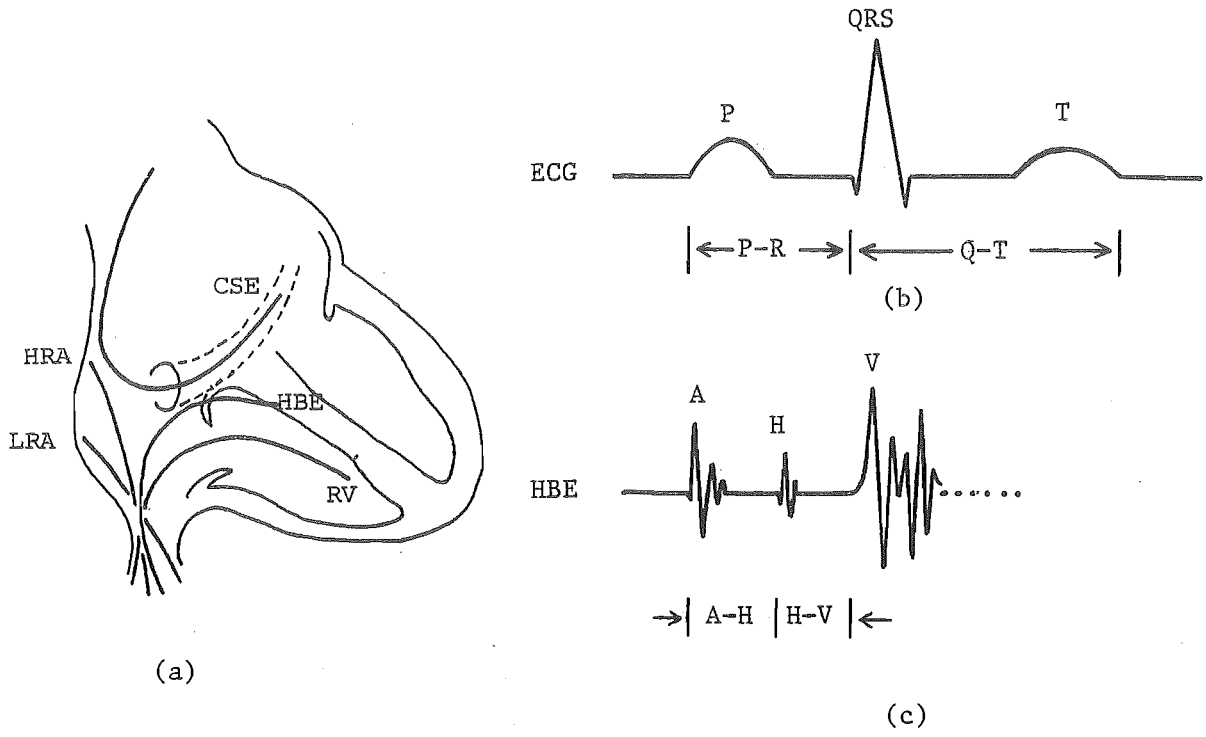


Fig. 4.4(a) Catheter electrode positions normally used for intracardiac recording. HRA = high right atrium, LRA = low right atrium, HBE = His bundle electrode, RV = right ventricle, and CSE = coronary sinus electrode. (From Curry, 1975, p. 48.)
(b) Form of the ECG measured on lead II.
(c) Form of a typical recording from the HBE.

taken for depolarisation of the ventricular myocardium. The Q-T interval (see Fig. 4.4(b)) has a normal variation of 300-420 ms and the T wave is of lower amplitude and longer duration than the QRS wave.

4.5 INTRACARDIAC ELECTROCARDIOGRAPHY

The ECG recorded on the body surface contains only information relating to activity of the atrial and ventricular myocardium. However, in order to assist in an accurate diagnosis of arrhythmias, it is often necessary to be able to detect activation of specific sites of the CCS which are not identifiable in the surface ECG. A number of workers, for example Berbari *et al.* (1979) and Bones (1981), have detected signals from the VSCS in the surface ECG using high gain, low noise circuitry and signal averaging techniques. However difficulties caused by measurement noise and the need for accurate time registration of the signals have so far prevented clinically useful interpretations of these signals from being made.

"Intracardiac electrocardiography" (also called "electrophysiological studies" or simply EP studies) is a technique for directly recording activation of various parts of the CCS tissue. The technique involves the positioning of one or more electrode catheters within the heart. The method is invasive (and, therefore, potentially dangerous) and expensive, hence it is used only when necessary and is not usually repeated in the same patient. Catheters are normally positioned only in the right-sided chambers of the heart so that access may be gained via peripheral veins. An important advantage of intracardiac electrocardiography is that an electrode catheter may be used to pace the heart at specific locations with a variety of pulse sequences. The response of the heart to these pacing sequences provides valuable information on the conduction properties of the CCS tissue.. The positions commonly used for catheters are shown in Fig. 4.4(a). The high right atrial (HRA) lead is positioned in the vicinity of the SAN to record early activation of the atria. Signals from the lower right atrial (LRA) and coronary sinus electrode (CSE) positions detect activation in the right and left sides of the atrial myocardium respectively. The His bundle electrode (HBE) is positioned near the bundle of His. The HBE is important as it is used to detect when activation exits from the AVN. The sort of signal obtained from the HBE is shown in Fig. 4.4(c). The symbols A, H and V are often used to indicate the atrial, bundle of His and ventricular activations respectively as recorded by the HBE as shown in Fig. 4.4(c). The A-H interval shown in Fig. 4.4(c) gives a measure of the conduction time through the AVN. The right ventricular (RV) electrode is positioned in the apex of the right ventricle.

An EP study involves the measurement of conduction times through various parts of the CCS during the patient's normal rhythm and in response to various sequences of pacing stimuli applied at various cardiac sites. EP study protocols are reviewed comprehensively by Curry (1975). Only the measurements relevant to the discussion in chapter 5 are described here. Because of the importance of the AVN and the ease of making HB recordings, many of the measurements are made only on the AVN.

The heart rate and the A-H and H-V intervals (see Fig. 4.4(c)) during the patient's normal rhythm are measured. The shapes of the surface lead P and QRS waves are monitored during the EP study as these may indicate the manner in which the atrial and ventricular myocardium is activated. The heart is then paced, at the atria, at rates between about 100 and 200

beats per minute (BPM). Conduction times through the AVN at the different rates are measured and a pacing curve of A-H versus A-A (often called the cycle length (CL)) is plotted - see Figs 4.5(a) and 4.6(a). The longest CL at which Wenckebach and 2:1 block occur are also measured.

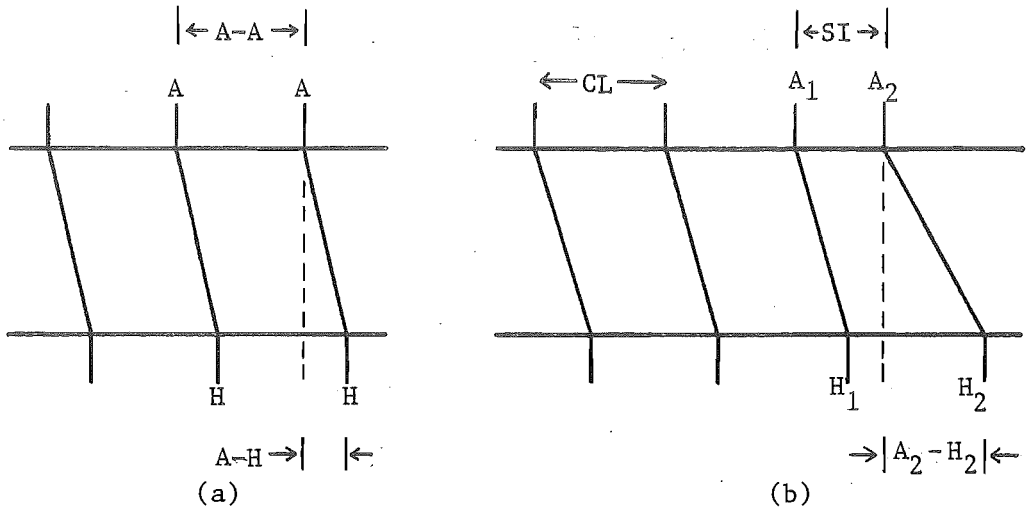


Fig. 4.5 Ladder diagrams of (a) constant pacing and (b) the extrastimulus test.

The refractory properties of the AVN are measured as follows using the extrastimulus technique. A number (at least eight) of pacing stimuli are delivered at a constant rate to the atria and these are followed by a single early stimulus (the "extrastimulus") at a period called the stimulus interval (SI) as shown in Fig. 4.5(b). The last regular stimulus and the early stimulus are labelled 1 and 2 respectively as shown in Fig. 4.5(b). The refractory properties (at this CL) are characterised by plotting a curve of either A_2-H_2 versus A_1-A_2 (called the refractory curve) or H_1-H_2 versus A_1-A_2 as shown in Fig. 4.6. These two curves both contain equivalent information. These curves are often summarised by two parameters called the effective refractory period (ERP) and the functional refractory period (FRP) which are indicated in Fig. 4.6. The ERP is the shortest SI for which the extrastimulus is conducted to the HB, and the FRP is the shortest interval which is measured at the HB. A number of authors (for example, Simpson *et al.*, 1979) have pointed out that the FRP is a complicated function of refractoriness and is not easily interpreted in terms of delay in conduction. However the FRP is useful in that it describes the minimum SI which can be coupled to following tissue. Usually extrastimulus tests are carried out at only one, or perhaps two, CL.

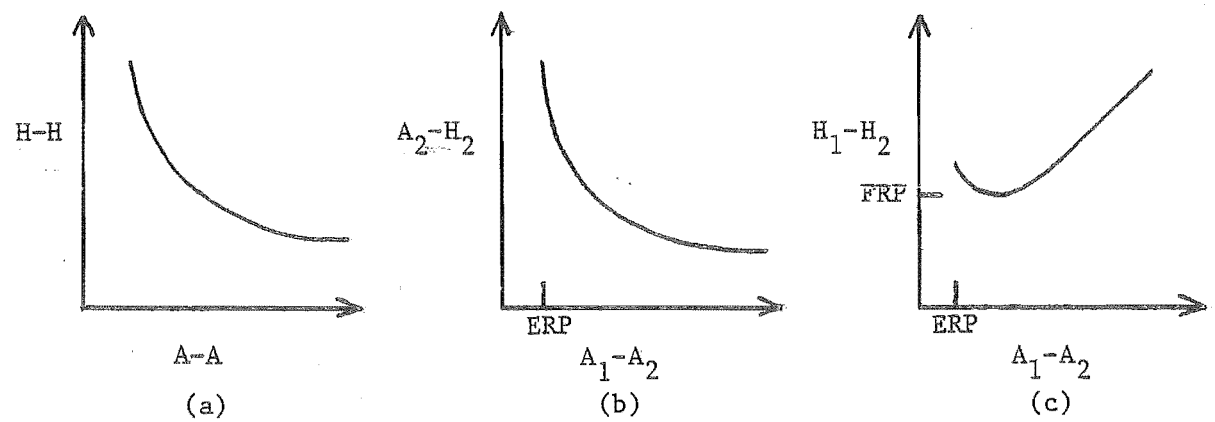


Fig. 4.6 Conduction curves for the AVN.
(a) Pacing curve.
(b) Refractory curve (obtained from extrastimulus data).
(c) H_1-H_2 versus A_1-A_2 curve (obtained from extrastimulus data).

If the patient suffers from an intermittent tachycardia, then the tachycardia may be triggered, or "precipitated", by some of the constant pacing or extrastimulus sequences described above. In fact, in such a patient, one attempts to trigger the tachycardia in this manner in order to learn something of its mechanism. Attempts are often made to terminate a tachycardia by pacing at a rate higher than the tachycardia rate. This is called overpacing. Some artificial pacemakers use overpacing to terminate tachycardias. Depending on the type of arrhythmia present, the pacing and extrastimulus tests may also be carried out at other sites such as the coronary sinus and the right ventricle in order to measure both the antegrade and retrograde conduction properties of different parts of the CCS. Some or all of the above tests may also be carried out after the administration of anti-arrhythmic drugs to assess the suitability of the drugs as treatment.

Information from the EP study, together with the surface ECG and other clinical data, is used by the cardiologist to try to determine the mechanism and physiological basis of a patient's arrhythmia. Various appropriate treatments may then be undertaken. However, the complexity of cardiac conduction and the relatively small amount of data collected during an EP study often make the precise determination of the mechanism of an arrhythmia extremely difficult. In chapter 5 a model of the CCS is described which may be used as an aid in arrhythmia diagnosis.

5. MODELLING CARDIAC ARRHYTHMIAS

5.1 INTRODUCTION

The diagnosis of cardiac arrhythmias is an inverse problem (as described in §1.1) since some of the system (heart) properties (electrophysiological parameters which are related to the presence of tissue disease or defects) are to be determined from the response (the EP data, the ECG and other clinical data) of the system to certain controlled inputs (pacing sequences, drugs etc.). Because the CCS is so complex, and the number of measurements available is relatively small, the determination of the cause of an arrhythmia directly from the clinical data is often not possible. The cardiologist draws on a large amount of empirical knowledge in an attempt to determine conduction sequences which are consistent with all the data. These conduction sequences may suggest possible underlying physiological defects. However the variability of cardiac conduction (particularly through the AVN) under different conditions, makes it difficult for him to accurately determine possible conduction sequences which are quantitatively consistent with the EP data.

As described in §1.1, inverse problems for complex systems can often only be usefully solved by model fitting procedures, especially if the number of measurements available is small. The model provides a convenient way of including the available *a priori* information on the system. Some of the unknown system parameters may be estimated directly from the measurements and others may be estimated by requiring that, in response to the same stimulus, the model outputs mimic the system outputs. The model parameters are then assumed to be reasonable estimates at the system parameters. This chapter describes a modelling approach to the inverse problem of cardiac arrhythmia diagnosis with a view to assisting the cardiologist in the interpretation of patients' EP data. This is facilitated by the construction of a model of the CCS. Data obtained from EP studies performed on a patient are used to estimate some of the model parameters. The model can then be used to investigate possible mechanisms for a patient's arrhythmia. The model may also be used to predict the response of a patient to a variety of therapies such as drugs and pacemakers.

Various ways of modelling cardiac conduction have already been reported. Miller (1979) describes numerical simulations of conduction, delay and block in cardiac tissue fibres based on the Hodgkin-Huxley equations (Hodgkin and Huxley, 1952; Noble, 1962). Networks of coupled excitable elements have been used as a basis for simulating action potential propagation in sheets of cardiac muscle tissue (Joyner *et al.*, 1975; van Capelle and Durrer, 1980) and in the AVN (Flanigan and Swain, 1967; Heethaar *et al.*, 1973c). These models are useful for simulating conduction in isolated sections of cardiac tissue but are not suitable for modelling the CCS as a whole, since it is very inhomogeneous and anisotropic. To model the CCS and relate the model to EP data, modelling at a functional rather than cellular level is required. Bhereur *et al.* (1971) describe a simple functional model consisting of two coupled relaxation oscillators representing the SAN and AVN. This model is able to simulate simple rhythms such as those that occur in second degree block but is not suitable (because of its simplicity) for quantitative simulation of arrhythmias in specific patients.

A CCS model (whose parameters are directly related to EP data) suitable for simulating arrhythmias, and which is relatively easy for a cardiologist to interact with, is described in §5.2. In §5.3 the implementation of the model and the output display format are described. Methods of estimating the model parameters from data obtained during routine patient EP studies are outlined in §5.4. Examples of simulations of two patients' arrhythmias are presented in §5.5. The significance of this work and potential applications of the model are described in §5.6.

Early work on this model has been published by Millane *et al.* (1980).

5.2 MODELLING CARDIAC CONDUCTION

The CCS model is made up of 11 sections (see Fig. 5.1), each representing a functional section of the CCS tissue. Ten of these sections make up the normal conduction system and the "bypass" section can be connected between any two of these sections to represent an accessory pathway. Each section has two ports (except the SAN which has one) through which impulses can enter and exit. Sections are connected by their ports so that instantaneous conduction can occur between connected sections. To allow the model to be used to represent a particular patient's conduction system, the electrophysiological properties of each section are described by a

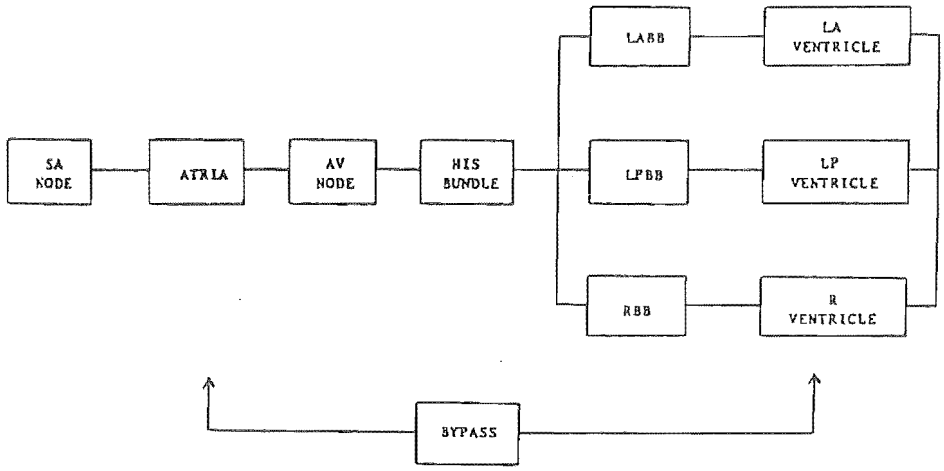


Fig. 5.1 Block diagram of the conduction system model.

number of variable parameters which are listed in table 5.1. The manner in which conduction through the sections has been modelled is described in terms of these parameters.

Table 5.1 Model parameters

Parameter*	Description
T	spontaneous period
c_o^A	basal conduction time (antegrade)
c_o^R	basal conduction time (retrograde)
ℓ_o	basal CL
e_o^A	basal ERP (antegrade)
e_o^R	basal ERP (retrograde)
d	maximum delay
α	(refractory curve time constant)$^{-1}$
β	rate of change of ERP with CL
γ	(CL time constant)$^{-1}$
b	effect of a concealed stimulus on the subsequent stimulus
τ_{cc}	length of concealed conduction zone

*Each parameter may be different for each section except that ℓ_o is the same for all sections and b and τ_{cc} apply only to the AVN.

The parameter T is the spontaneous period and represents the pacemaker properties of a section. If a section does not receive a stimulus within a time interval T, then a spontaneous impulse is generated in the centre of the section and conducts in both directions.

The conduction time through a section from one port to the other, denoted by $c = c(\ell,s)$, depends on the CL, ℓ, and SI, s, associated with the stimulus when it arrives at the section. Conduction times through tissue other than the AVN are almost independent of CL and SI (except, sometimes, for SI very close to the ERP) (Ferrier and Dressel, 1973) and so conduction times would usually be set to constant values in the non-AVN sections. Experimental studies on the rat AVN by Heethaar *et al.* (1973b) and the human AVN by Teague *et al.* (1976) show that the conduction time (for a particular CL) can be reasonably well approximated by an exponential function of the SI, so the modelled conduction time is defined by

$$\left.\begin{aligned} c(\ell,s) &= \infty && , \quad s < e \\ &= c_o + d\,\exp(-\alpha(s-e)) && , \quad s > e \end{aligned}\right\} \quad (5.1)$$

where $c = \infty$ means that the stimulus is blocked and d, α and e are the maximum delay, (time constant)$^{-1}$ and ERP respectively. The parameter c_o is called the basal conduction time and is equal to the conduction time in the limit of long SI. If the delay drops to 5% of its maximum value for a SI which is 300 ms longer than the ERP, then $\alpha = 0.01$ ms^{-1}. Equation (5.1) defines the modelled refractory curve described in §4.5.

In general, the shape of the refractory curve varies with CL in an ill-defined manner. In fact no comprehensive study of the dependence of refractory curves on CL appears in the literature. Since very little EP data which allows one to characterise the change in the refractory curve with CL is routinely available, it is pointless to characterise it by more than one parameter. It seems reasonable to expect c_o to be independent of CL since the CL would have no effect after a sufficiently long SI. It is well known (see, for example, Denes *et al.*, 1974) that the refractoriness of all cardiac tissue depends on the CL at which it is measured. Hence the refractory curve is made a function of CL by making the ERP depend on CL. Denes *et al.* (1974) show that the ERP of both the AVN and atrial tissue varies approximately linearly with CL for CL less than about 800 ms. As would be expected, the refractoriness of cardiac tissue tends to approach a constant value for sufficiently long CL (Janse *et al.*, 1969). Based on

this information, the ERP, $e = e(\ell)$, is modelled as

$$\left.\begin{aligned} e(\ell) &= e_o + \beta(\ell - \ell_o) \quad , \quad \ell < \ell_o \\ &= e_o \quad , \quad \ell > \ell_o \quad , \end{aligned}\right\} \quad (5.2)$$

where the constants e_o and ℓ_o are called the basal ERP and basal CL respectively, and β is a measure of the rate of change of ERP with CL. It is important to note that $\beta < 0$ for the AVN whereas $\beta > 0$ for other cardiac tissue (Denes *et al.*, 1974). This means that the AVN adapts to block the conduction of high stimulus rates whereas the other tissue adapts to allow conduction of higher rates. This difference between AVN and non-AVN tissue is discussed with reference to examples in §5.5. The minimum constant pacing period $s_{1:1}$ which produces 1:1 conduction, occurs when $e(s_{1:1}) = s_{1:1}$, so that, using (5.2),

$$s_{1:1} = (e_o - \beta\ell_o)/(1 - \beta) \ .$$

Inspection of (5.1) and (5.2) shows that the modelled conduction time is defined by

$$\left.\begin{aligned} c(\ell,s) &= \infty \quad , \quad s < e \\ &= c_o + d \exp[-\alpha(s - e_o - \beta(\ell - \ell_o))] \quad , \quad \ell < \ell_o \\ &= c_o + d \exp[-\alpha(s - e_o)] \quad , \quad \ell > \ell_o \end{aligned}\right\} s > e \quad (5.3)$$

Especially under pathological conditions, conduction and refractoriness properties may be different in the antegrade and retrograde directions (cf. Akhtar *et al.*, 1975). Since little information on retrograde conduction is usually obtained during routine EP studies, only the model parameters c_o and e_o can be different for the two directions of conduction. The symbols c_o^A, e_o^A, c_o^R and e_o^R indicate the antegrade and retrograde values respectively of these parameters.

It is worth noting that the variation of refractoriness with CL described by (5.2) only strictly applies while the AVN is being externally paced. During sinus rhythm the autonomic nervous system causes AVN refractoriness to vary with CL in the opposite manner to that during pacing (Curry, 1975, p. 54). The significance of this is further discussed in §6.3.

The definition of CL in §4.5 applies to the extrastimulus situation - i.e. where the previous stimulus is preceded by a sequence of equally spaced stimuli. However the conduction times and ERP in the model must

be defined for any stimulus sequence that may occur. A number of investigators, such as Janse *et al.* (1969), Han and Moe (1969) and Heethaar *et al.* (1973a), have examined the conduction properties of the AVN and other cardiac tissue in response to step changes in stimulation rate. Studies of this type show that refractoriness continues to change significantly for approximately five seconds following the change of rate. This has sometimes been referred to as the "cumulative effect of CL" and simply means that the tissue adapts slowly to changes in stimulation rate. This effect is modelled by defining the CL, ℓ_n, after a stimulus of SI s by

$$\ell_n = \ell_{n-1} \exp(-\gamma s) + s(1 - \exp(-\gamma s)) \qquad (5.4)$$

where ℓ_{n-1} is the CL before the stimulus and γ is a (time constant)$^{-1}$. Heethaar *et al.* (1973a) model conduction time in the AVN in a similar manner to the definition of ℓ_n in (5.4). Applying (5.4) to the case of a step change in stimulation rate from period ℓ to ℓ', then the CL, ℓ_n, n stimuli after the change is given by

$$\ell_n = \ell' + (\ell - \ell') \exp(-n\gamma\ell') \ . \qquad (5.5)$$

Equation (5.5) shows that the conduction properties of the AVN following a step change in stimulation rate are an exponential function of time (equal to $n\ell'$) following the change, in agreement with the observations of Heethaar *et al.* (1973a). If the relative change of ℓ_n from ℓ to ℓ' is 95% after 2 seconds, then $\gamma = 0.0015$ ms^{-1}.

If a stimulus arrives at a section with a SI less than the ERP, then it is blocked and does not enter the section. The manner in which a following stimulus conducts through the section, depends on the effect that the blocked stimulus has on the tissue's refractoriness. The effect of a blocked stimulus on the conduction time of a following impulse in the AVN has been examined by Langendorf (1948) and Wu *et al.* (1976), and the results of their work are summarised as follows. To avoid confusing notation, the time interval between stimuli A_i and A_j (denoted by A_i-A_j in §4.5) is denoted by s_{ij} in this section. Consider the pacing sequence shown in Fig. 5.2 consisting of three stimuli labelled A_1, A_2 and A_3. A_1 is the last stimulus of a regular sequence whose period defines the CL ℓ, and s_{12} is less than the ERP, $e(\ell)$, so that A_2 is blocked. A_3 is a test stimulus whose conduction time A_3-H_3 is examined. It is found that if $s_{12} < e(\ell) - \tau_{cc}$, where τ_{cc} is called here the length of the concealed conduction zone (CCZ), then the conduction time A_3-H_3 is equal to what it would be if A_2 was not present. Hence the SI associated with A_3 is effectively s_{13}.

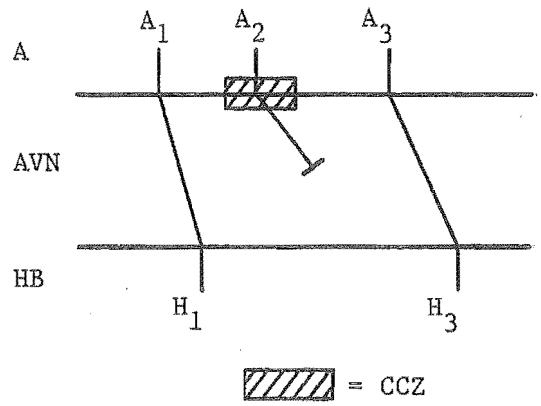


Fig. 5.2 Pacing sequence used by Wu *et al.* (1976) to examine concealed conduction.

This is interpreted as being because A_2 does not enter the AVN tissue (sometimes called "entrance block") and so has no effect on its refractoriness. However if $e(\ell) - \tau_{cc} < s_{12} < e(\ell)$ then it is found that the conduction time of A_3 is longer than it would be if A_2 was not present. Hence, in this case, A_2 has depressed the conduction of A_3. This is interpreted as being due to A_2 penetrating the AVN tissue (and so making the tissue more refractory) but the AP "dying out" before it exists from the AVN. This is often referred to as <u>concealed</u> or <u>decramental</u> conduction. The interval $(e(\ell) - \tau_{cc}, e(\ell))$ is called the CCZ. Concealed conduction is modelled by reducing the SI associated with a stimulus, if the previous stimulus falls in the CCZ. Since few, if any, measurements obtained during routine EP studies can be used to characterise concealed conduction, it is pointless to use more than one parameter (in addition to τ_{cc}) to quantify concealed conduction in the model. The modelled SI, s_3, of the third stimulus, A_3, of a consecutive sequence of stimuli A_1, A_2, A_3 is defined by

$$\left.\begin{aligned} s_3 &= s_{13} && , \quad s_{12} < e_2 - \tau_{cc} \\ &= b\, s_{12} + s_{23} && , \quad e_2 - \tau_{cc} < s_{12} < e_2 \\ &= s_{23} && , \quad s_{12} > e_2 \end{aligned}\right\} \quad (5.6)$$

where e_2 is the ERP associated with A_2 and b is a constant such that $0 \leq b \leq 1$. The parameter b reflects the degree to which the concealed stimulus affects conduction of the following stimulus. If $b = 1$ then the concealed stimulus does not affect the following stimulus. The definition of s_3 in (5.6) is consistent with the observations of Wu *et al.* (1976), that a large value of s_{12} increases the conduction time of A_3 more than

does a small value. Measurement of the concealed conduction zone is hampered by the atrial FRP limiting the minimum SI which can be coupled to the AVN. The results of Wu *et al.* (1976) suggest that on average $\tau_{cc} > 100$ ms. Inspection of the data for case 8 given in Wu *et al.* (1976) (using values extracted from their Figure 4 for A_3-H_3 = 230 ms) and making use of (5.6) gives a rough estimate for b of 0.5 for this patient.

Nominal values for the model parameters are given in table 5.2. It is instructive to examine the way in which the model simulates Wenckebach block.

Table 5.2 Nominal values for model parameters.

Section	T (ms)	ℓ_o (ms)	c_o^A and c_o^R (ms)	e_o^A and e_o^R (ms)	d (ms)	α (ms^{-1})	β	γ (ms^{-1})	b	τ_{cc} (ms)
SAN	800	↑	5	200	0	0	0	0.0015	-	-
Atria	1200		40	200	0	0	0.15	0.0015	-	-
AVN	1000		80	300	120	0.01	-0.18	0.0015	0.5	100
HB	1300	1000	10	200	0	0	0.15	0.0015	-	-
3 bundle branches	1500		20	200	0	0	0.15	0.0015	-	-
3 ventricles	1800		40	200	0	0	0.15	0.0015	-	-
Bypass	1800	↓	20	200	0	0	0.15	0.0015	-	-

Consider the stimulation, with period s, of the AVN producing 4:3 Wenckebach as shown in Fig. 5.3(a). The model refractory curves seen by each stimulus in the cycle are shown in Fig. 5.3(b). The SI of A_1 is long (equal to $(1+b)s$ or $2s$, depending on whether or not A_4 falls in the CCZ) and so A_1 is conducted rapidly. This long SI causes the CL to lengthen and so the AVN becomes less refractory (since $\beta < 0$ for the AVN) allowing A_2 to be conducted. Following A_2, the CL begins to shorten because of the small SI and the AVN becomes increasingly refractory (since $\beta < 0$) so that subsequent impulses are conducted with increasing delay until stimulus A_4 is blocked. The cycle then repeats. Note that this mechanism requires that $\beta < 0$ so that Wenckebach is usually observed only in the AVN. The mechanism for Wenckebach described above is consistent with that of Simpson *et al.* (1978) who measured

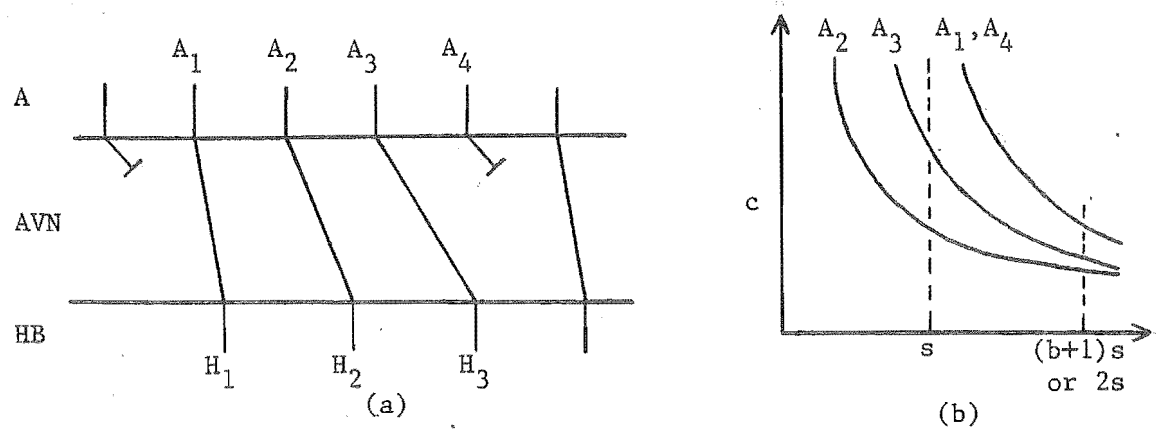


Fig. 5.3 The mechanism of Wenckebach cycles.
(a) Ladder diagram of 4:3 Wenckebach.
(b) Refractory curves associated with each of the four stimuli.

(using rather sophisticated programmed pacing) refractory curves following each stimulus during Wenckebach. However they interpret the SI of A_1 as being s, and A_4 producing a decrease in refractoriness.

5.3 IMPLEMENTATION OF THE MODEL

The conduction system model described in §5.2 was implemented as an interactive program on a PDP 11/70 computer, interaction being via an alphanumeric VDU. On running the program, the operator first connects the bypass section (if it is required) between the required sections. The parameters of each section are displayed on the VDU and the operator can change any of these to the values he requires. A facility is provided for inserting stimuli in any section and at any time during the simulation to represent pacing sequences or ectopics. The operator specifies any such stimuli that he requires and labels them as "paced" or "ectopic". The simulation time is entered by the operator and the simulation runs.

At the end of the simulation the operator can have a ladder diagram and an ECG trace displayed. An example of a ladder diagram produced by the model is shown in Fig. 5.4. The model sections are listed vertically and time increases from left to right. The characters printed on the ladder diagram indicate the times at which the sections are activated. To aid interpretation, six characters are used to show the type of impulse These characters are A, R, S, P, E and B which indicate antegrade, retrograde, spontaneous, paced, ectopic and a stimulus conducted from the bypass respectively. The ladder diagram is displayed to the operator on the VDU and may be printed out on the line printer.

NO BYPASS

PACING AND ECTOPIC SEQUENCE

SI	N	SEC	TYPE
360	30	2	P

SIMULATION TIME = 10 SEC

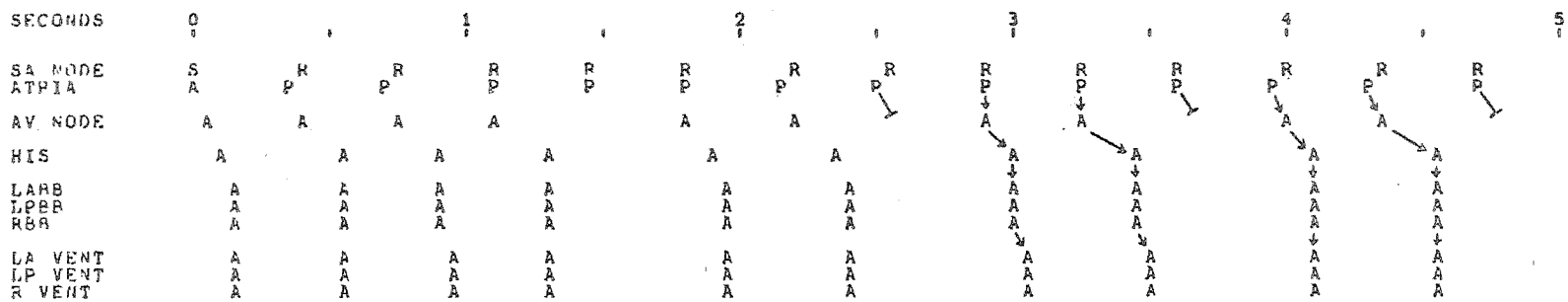


Fig. 5.4 Simulation of a "normal heart" paced at a CL of 360 ms showing the development of 3:2 AVN Wenckebach.

The ECG trace is displayed on a graphics terminal, and an example is shown in Fig. 5.4. In the examples presented in this chapter, the ECG trace represents lead II (see §4.4). The ECG, E(t), is made up of waveforms that are initiated each time the atrial or one of the ventricular sections is activated. Hence the ECG is defined by

$$E(t) = \sum_i f_i(t - t_i)$$

where $f_i(t)$ is the waveform of the section activated at time $t = t_i$, if the section was activated at time $t = 0$. The waveforms $f_i(t)$ (for lead II) in are shown in table 5.3 and have been chosen so that they produce (as closely as possible for such a simple model of the ECG) typical bundle branch block patterns (refer to Goldman, 1979, ch. 9) when these forms of block occur. As indicated in table 5.3, the waveforms depend on the manner in which the sections are activated. Because of the simplified representation of the atrial and ventricular myocardium used in the model, the ECG provides accurate timing information but the shape of the ECG does not provide reliable clinical information.

Table 5.3 ECG waveforms (lead II)
(scale: ⊢—⊣ = 50 ms).

Section	Type of Impulse		
	A	R	S, P, E, B
Atria			
LA Ventricle			
LP Ventricle			
R Ventricle			

Fig. 5.5 ECG traces of a "normal heart" paced at different rates (see text).

If he wishes, the operator can also obtain a detailed printout which shows the simulation time, type of stimulus, CL, SI, ERP, and conduction time, each time a section is stimulated.

Fig. 5.4 shows the output from the model when it is used to simulate a "normal heart" (i.e. using the parameters in table 5.2) paced at the atria at a CL of 360 ms. Arrows have been added to the ladder diagrams shown in this chapter, to clarify the conduction sequences. After 1.5 seconds from the beginning of the simulation, the AVN develops stable 3:2 Wenckebach block. The ECG shows the characteristic increasing P-R interval following the block stimulus. Fig. 5.5 shows ECG traces for a simulation of a normal heart paced at various rates at the atria. The AVN conducts normally at a CL of 800 ms, suffers first degree block at 400 ms, Wenckebach (4:3 at 370 ms and 3:2 at 360 ms) and type two second degree block (2:1) at 345 ms, simulating the increasing degree of block with increasing pacing rate described as in §4.3. The bottom trace in Fig. 5.5 shows a simulation of complete AV block. The ventricles are activated by spontaneous depolarisation in the AVCS, and there is complete dissociation of the P and QRS waves.

5.4 PROCESSING MEASURED ELECTROPHYSIOLOGICAL DATA

In order to use the CCS model to simulate patients' arrhythmias, the model parameters described in §5.2 must be estimated from the measured EP data. As mentioned in §5.2, some of these measurements will usually only apply to the AVN. The values of T for the SAN and c_o for all the sections can be obtained directly from recordings made with the patient in sinus rhythm. It is assumed that extrastimulus measurements are made at only one CL (which is often the case), denoted here by ℓ_e, and the subscript e is used to denote quantities measured at $\ell = \ell_e$. Hence e_e is obtained directly from the measured refractory curve. By fitting a straight line (using least squares) to the extrastimulus data in the form $\ln(c - c_o)$ versus $(s - e_e)$, d and α can be estimated.

Evaluating (5.3) for constant pacing at a period s (i.e. $\ell = s$), shows that

$$c(s,s) = c_o + d \exp[-\alpha(s - e_o - \beta(s - \ell_o))] \quad , \quad s < \ell_o \ , \tag{5.7}$$

where c(s,s) is the conduction time for constant pacing, and evaluating

(5.2) at $\ell = \ell_e$ gives

$$e_e = e_o + \beta(\ell_e - \ell_o) \quad , \quad \ell_e < \ell_o \ . \tag{5.8}$$

Making use of (5.7) and (5.8) shows that

$$c(s,s) = c_o + d \exp[-\alpha(\beta\ell_e - e_e + (1-\beta)s)] \quad , \quad s < \ell_o \ . \tag{5.9}$$

Since, usually, $s < \ell_o$ for the constant pacing measurements, fitting a straight line to the pacing data in the form of $\ln(c(s,s) - c_o)$ versus s and using (5.9) gives an estimate of β. A reasonable value (say 1000 ms) is chosen for ℓ_o, and e_o is determined using (5.8).

It is usually possible to pace the heart such that the AVN exhibits stable 2:1 block. Since 2:1 block occurs at a SI which is close to the ERP, the blocked stimulus is assumed to undergo concealed conduction. Let $s_{2:1}$ and $c_{2:1}$ be the pacing period at which 2:1 conduction is achieved and the measured conduction time of the conducted stimulus during 2:1 conduction respectively. Applying (5.3) and (5.6) to stable 2:1 conduction shows that

$$c_{2:1} = c_o + d \exp[-\alpha((b+1)s_{2:1} - e_o - \beta((b+1)s_{2:1} - \ell_o))] \tag{5.10}$$

which on rearrangement gives

$$b = [\ln(d/(c_{2:1} - c_o)) - \alpha(\beta\ell_o - e_o)]/[(1-\beta)\,\alpha\, s_{2:1}] - 1 \ . \tag{5.11}$$

Hence (5.11) can be used to estimate b from EP measurements made during stable 2:1 conduction. In order that b is consistent with 2:1 conduction at this SI the conditions

$$\left.\begin{aligned} & s_{2:1} < e(s_{2:1}) \\ \text{and} \quad & \\ & (b+1)s_{2:1} > e((b+1)s_{2:1}) \end{aligned}\right\} \tag{5.12}$$

must be satisfied. If (5.12) is not satisfied then b must be adjusted until (5.12) is satisfied and $c_{2:1}$ given by (5.10) is reasonably close to that measured.

In order to estimate γ, measurement of conduction times during a period of varying CL is required. The most convenient way to estimate γ would be to measure the conduction times of stimuli following a step change in stimulation rate. Then γ could be estimated using (5.3) and (5.5).

However measurements of this type are not usually made during routine EP studies. Another situation which involves a varying CL is Wenckebach, as described in §5.2. Measurements made during stable 3:2 Wenckebach may be used to estimate γ as follows. Because Wenckebach occurs at a SI close to the ERP, the blocked stimuli are assumed to fall within the CCZ. Let s_W be the SI at which stable 3:2 Wenckebach is achieved and let c_{W1} and c_{W2} be the measured conduction times of the stimuli A_1 and A_2 respectively as shown in Fig. 5.6.

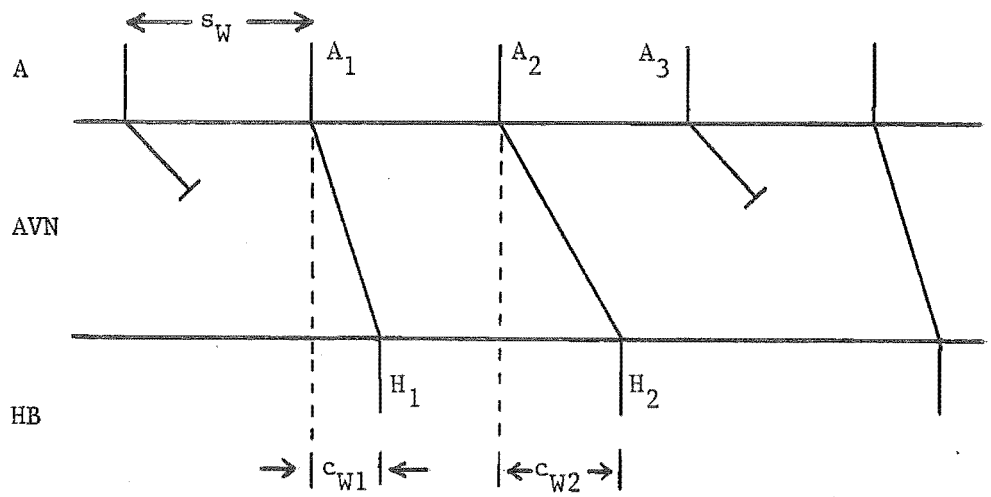


Fig. 5.6 Illustration of EP quantities defined for 3:2 Wenckebach.

Let ℓ_{W1} and ℓ_{W2} be the CL associated with the stimuli A_1 and A_2 respectively. The values of ℓ_{W1} and ℓ_{W2} can be computed from the measurements s_W, c_{W1} and c_{W2} using (5.3) and (5.6). Making use of (5.4) and (5.6) gives

$$\ell_{W1} = \ell_{W2} \exp(-\gamma s_W) + s_W (1 - \exp(-\gamma s_W)) \tag{5.13}$$

and

$$\ell_{W2} = \ell_{W1} \exp(-\gamma(b+1)s_W) + (b+1)\, s_W (1 - \exp(-\gamma(b+1)s_W)) \ . \tag{5.14}$$

Inspection of (5.13) shows that γ can be estimated as

$$\gamma = (1/s_W) \ln((\ell_{W2} - s_W)/(\ell_{W1} - s_W)) \ . \tag{5.15}$$

In order that 3:2 Wenckebach occurs at a SI of s_W, the conditions

$$\left.\begin{aligned} & s_W < e(\ell_{W1}) \\ & (b+1)s_W > e(\ell_{W1}) \\ \text{and} \quad & s_W > e(\ell_{W2}) \end{aligned}\right\} \tag{5.16}$$

must be satisfied. Adjustments may have to be made to b and γ to ensure that (5.16) is satisfied and also that (5.13) and (5.14) are reasonably well satisfied.

An interactive computer program was written which performs the calculations described above and this is used by the operator to assist in determining the model parameters from the patient's measured EP data. Sufficient EP data is not usually available to allow determination of all the model parameters so these must be estimated from either other data or an "educated guess".

5.5 EXAMPLES

The CCS model has been used to simulate arrhythmias in two patients who were examined at The Princess Margaret Hospital, Christchurch. Both of these patients suffered from an arrhythmia called the Wolff-Parkinson-White (WPW) syndrome (Krikler and Wellens, 1975). However they exhibited slightly different characteristics during EP studies and the model simulations suggest a possible physiological basis for these differences. The study did not involve experimentation on the patients as the EP data was obtained from existing records. The patients' EP data (in the form of chart recordings of catheter electrode traces) obtained during routine EP studies were examined and the EP parameters described in §4.5 determined. These were processed, as described in §5.4, to estimate the major CCS model parameters. The model was then used to simulate the patients' arrhythmias.

The WPW syndrome is described in §5.5.1, and the two patients' EP data, model parameters and simulations are outlined in §5.5.2 and §5.5.3. The results of the simulations are used to compare the two patients' arrhythmias in §5.5.4.

5.5.1 The Wolff-Parkinson-White Syndrome

The WPW syndrome is an example of one of the pre-excitation arrhythmias described in §4.3 and involves an accessory pathway between the atria and ventricles. In the classical form (which is that considered here) of the WPW syndrome, the accessory pathway forms a direct connection between the atria and ventricles. In other forms of the syndrome, the accessory pathway may involve the AVN or HB (Krikler and Wellens, 1975). The main ECG features of the WPW syndrome are a short P-R interval and a wide QRS

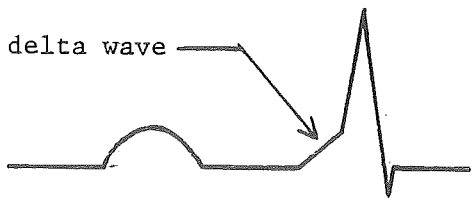


Fig. 5.7 Form of the ECG for the WPW syndrome.

wave as shown in Fig. 5.7. The sinus impulse is conducted through both the AVCS and the accessory pathway, however the accessory pathway conducts the impulse more rapidly as it does not possess the delaying properties of the AVN. Hence the ventricles begin to depolarise early, resulting in an early initiation of the QRS wave and a short P-R interval. Because the impulse that first reaches the ventricles has not travelled through the bundle branch system, depolarisation of the myocardium begins slowly and the initial part of the QRS is not as steep as usual - see Fig. 5.7. This early section of the QRS is called a delta wave. Once the impulse that conducts through the AVCS reaches the ventricles, depolarisation of the myocardium is rapid and the rest of the QRS is normal.

The WPW syndrome is not, in itself, dangerous, however patients with this syndrome often experience associated tachycardias. These tachycardias are usually of the "re-entrant" type (Wit and Cranefield, 1978) since they involve propagation of impulses around a closed circuit consisting of the AVCS, ventricular tissue, accessory pathway and atrial tissue. The tissue in this closed circuit must have specific conduction times and refractory periods for a tachycardia to occur. In addition, the tachcardia is usually only triggered by specific stimulus sequences. These tachycardias can be dangerous if they allow the ventricles to sustain high rates for significant periods of time. If the patient's CCS can be modelled reasonably accurately, experimentation with the model may allow an assessment of the patient's susceptibility to high ventricular rates.

5.5.2 Patient 1

Results of routine EP studies on patient 1 are listed in table 5.4. Episodes of tachycardia were precipitated by the extrastimulus testing for SI between 329 and 350 ms. The period of the tachycardia was 337 ms and it was sustained only for between two and six beats. The QRS during the tachycardia was similar to the normal QRS. Attempts to precipitate the tachycardia by constant pacing at 337 ms were unsuccessful.

Table 5.4 EP data for patient 1.

Sinus Rhythm:	CL = 1032 ms A-H = 80 ms H-V = 18 ms QRS has a delta wave
Atrial Pacing:	CL range = 300 to 600 ms A-H/A-A measurements: see Fig. 5.8 CL > 500 ms : QRS has a delta wave, H-V = 18 ms CL < 500 ms : QRS normalises (delta wave disappears) and H-V interval lengthens to 42 ms
Atrial Extrastimulus:	CL = 605 ms SI range = 300 to 500 ms A_2-H_2/A_1-A_2 measurements: see Fig. 5.8

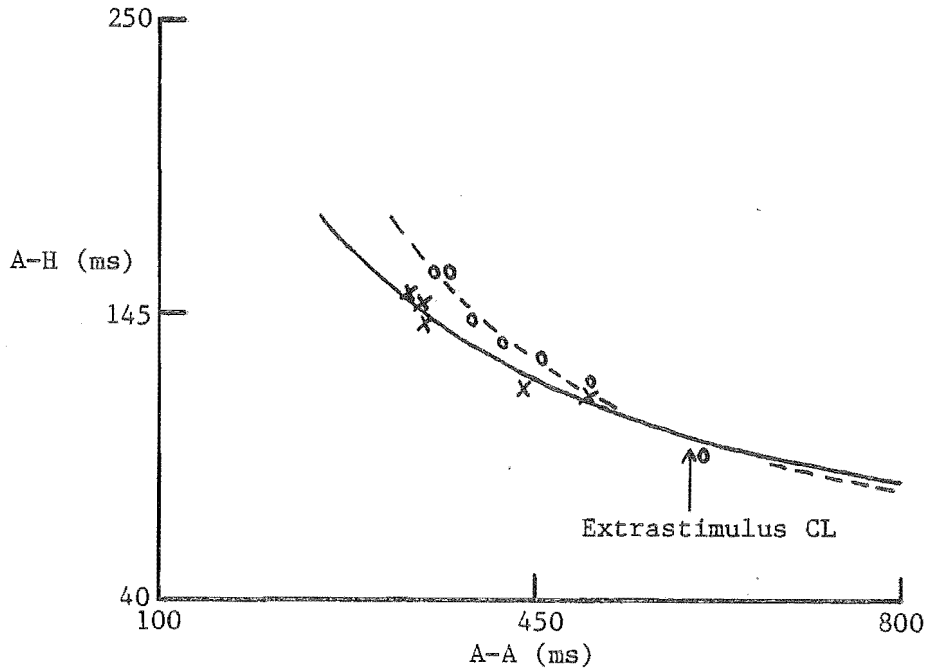


Fig. 5.8 Measured pacing (o) and extrastimulus (x) data, and modelled pacing (---) and refractory (——) curves for patient 1.

The above data indicates the WPW syndrome with early activation of the ventricles via antegrade conduction through the accessory pathway for a CL greater than 500 ms. For a CL less than 500 ms the accessory pathway is refractory and normal conduction through the AVCS occurs. The tachycardia could be triggered by an extrastimulus but not by constant pacing. Because the tachycardia has a normal QRS, conduction is antegrade through the AVCS and retrograde through the accessory pathway during the tachycardia.

The model parameters were estimated from the EP data as described in §5.4 and are listed in table 5.5. Fig. 5.8 shows the modelled pacing and refractory curves that have been fitted to the measured EP data. The bypass section is connected between the atria and the LA ventricle.

Table 5.5 Model parameters for patient 1.

Section	T (ms)	ℓ_o (ms)	c_o^A (ms)	c_o^R (ms)	e_o^A (ms)	e_o^R (ms)	d (ms)	α (ms^{-1})	β	γ (ms^{-1})	b	τ_{cc} (ms)
SAN	1032	↑	2	2	300	300	0	0	0	0.0015	-	-
Atria	2000	|	10	10	300	300	0	0	0.15	0.0015	-	-
AVN	2000	|	70	70	200	200	100	0.004	-0.24	0.0015	0.5	100
HB	2000	1000	10	10	300	300	0	0	0.15	0.0015	-	-
3 bundle branches	2000	|	32	32	300	300	0	0	0.15	0.0015	-	-
3 ventricles	2000	|	40	40	300	300	0	0	0.15	0.0015	-	-
Bypass	2000	↓	88	80	590	400	0	0	0.15	0.0015	-	-

Figs 5.9 to 5.11 show simulations of this patient's conduction sequences. Fig. 5.9 shows a simulation of the patient's sinus rhythm. The ladder diagram shows that the sinus impulse is conducted to the ventricles by the accessory pathway. The ECG shows a short P-R interval and a wide QRS with a delta wave. The result of constant pacing at the atria with six stimuli at the tachycardia period of 337 ms is shown in Fig. 5.10. At the end of the pacing sequence, the tachycardia is not triggered and the model reverts to sinus rhythm. Reference to the ladder diagram shows that a re-entrant circuit is not established because the AVN is refractory at this rate. Fig. 5.11 shows the simulated response to an extrastimulus test with CL = 605 ms and SI = 337 ms. The extrastimulus has initiated a tachycardia with a normal QRS. The period of the tachycardia is 342 ms. Inspection of the ladder diagram shows that after three beats, the tachycardia reverts to sinus rhythm because the AVN has adapted to the higher rate and has become refractory. Further simulations showed that at least one beat of a tachycardia is obtained for SI between 300 ms and 350 ms.

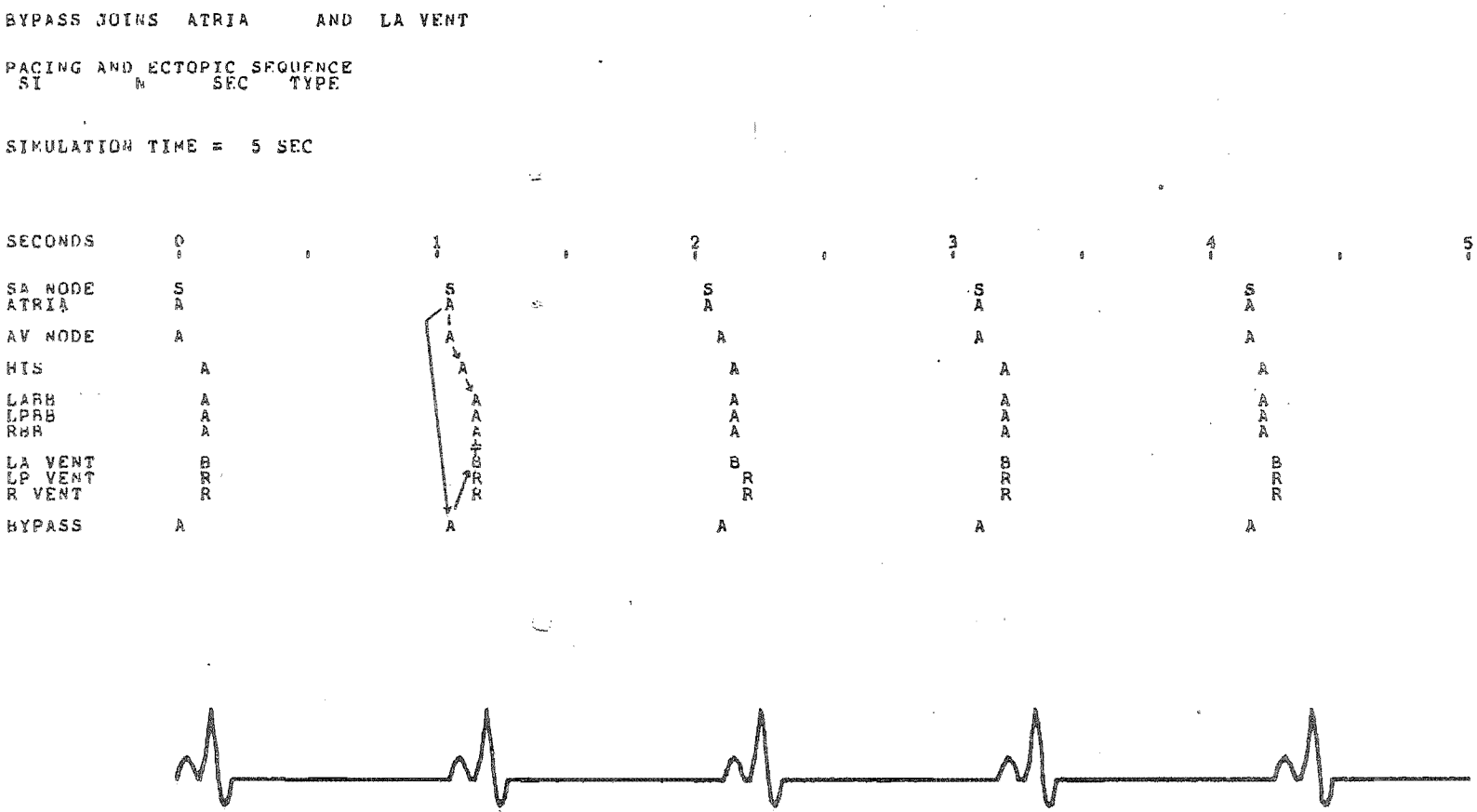


Fig. 5.9 Simulation of patient 1 in sinus rhythm (refer to text).

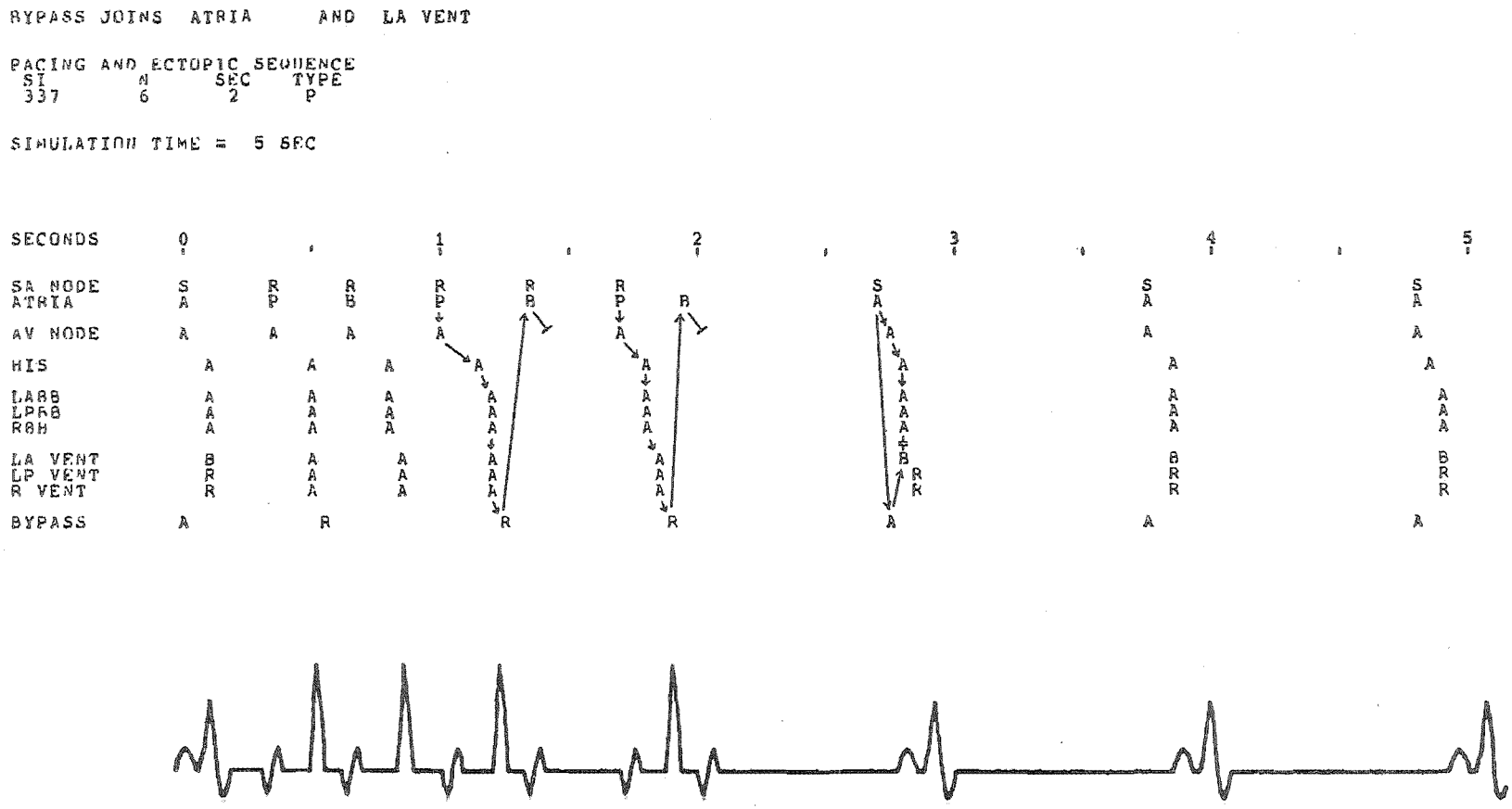


Fig. 5.10 Simulation of constant pacing at the atria for patient 1 (refer to text).

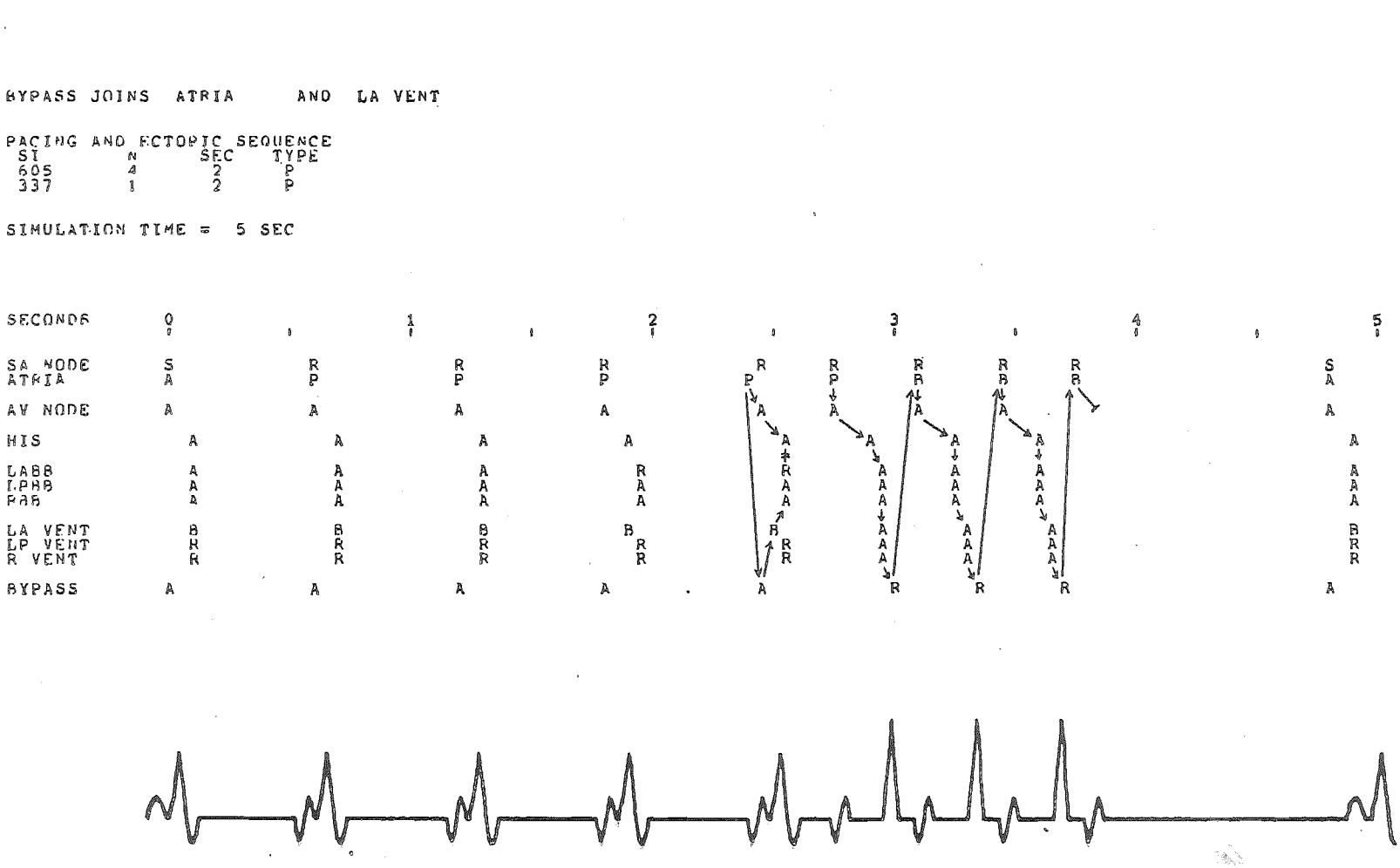


Fig. 5.11 Simulation of the initiation of patient 1's tachycardia by atrial extrastimulus testing (refer to text).

5.5.3 Patient 2

Routine EP study results for patient 2 are listed in table 5.6.

Table 5.6 EP data for patient 2.

Sinus Rhythm:	CL = 721 ms A-H = 75 ms H-V = 5 ms QRS has a delta wave
Atrial Pacing:	CL range = 300 to 600 ms CL > 360 ms : QRS has a delta wave, H-V = 5 ms CL < 360 ms : QRS normalises (delta wave disappears) and H-V interval lengthens to 20 ms
Atrial Extrastimulus:	CL = 569 ms SI range = 200 to 550 ms

Episodes of tachycardia were precipitated by constant pacing at a CL of 310 ms. The period of the tachycardia was 294 ms, it had a normal QRS and was stable (i.e. it did not spontaneously terminate). The tachycardia was not precipitated during the extrastimulus testing. The tachycardia could be terminated (sometimes) by pacing the atria at a CL of 284 ms.

The above data indicates the WPW syndrome of the same type as for patient 1. However this patient's tachycardia was triggered by constant pacing but not by an extrastimulus and was stable. The model parameters were estimated from the EP data and are listed in table 5.7. The bypass section is connected between the atria and the LA ventricle.

Table 5.7 Model parameters for patient 2.

Section	T (ms)	ℓ_o (ms)	c_o^A (ms)	c_o^R (ms)	e_o^A (ms)	e_o^R (ms)	d (ms)	α (ms^{-1})	β	γ (ms^{-1})	b	τ_{cc} (ms)
SAN	721	↑	2	2	250	250	0	0	0	0.0015	-	-
Atria	2000	|	10	10	300	300	0	0	0.15	0.0015	-	-
AVN	2000	|	60	60	150	150	126	0.008	-0.23	0.0015	0.5	100
HB	2000	800	5	5	300	300	0	0	0.15	0.0015	-	-
3 bundle branches	2000	|	15	15	300	300	0	0	0.15	0.0015	-	-
3 ventricles	2000	|	40	40	300	300	0	0	0.15	0.0015	-	-
Bypass	2000	↓	70	70	400	340	0	0	0.15	0.0015	-	-

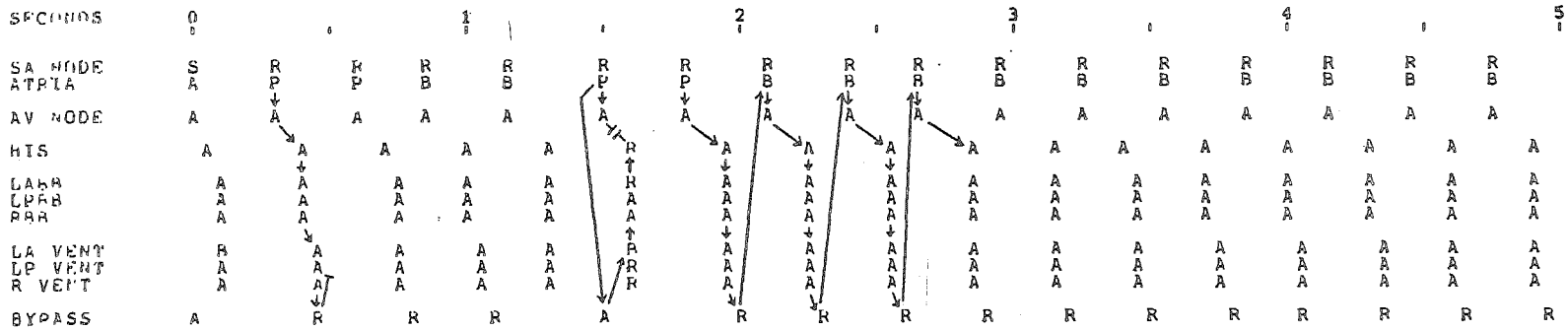
Figs 5.12 and 5.13 show simulations of the patient's rhythms.

BYPASS JOINS ATRIA AND LA VENT

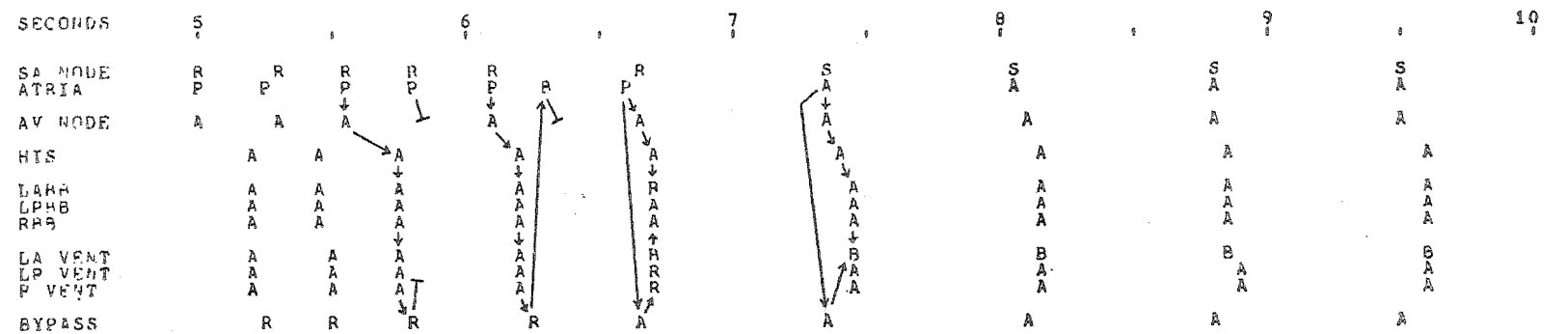
PACING AND ECTOPIC SEQUENCE

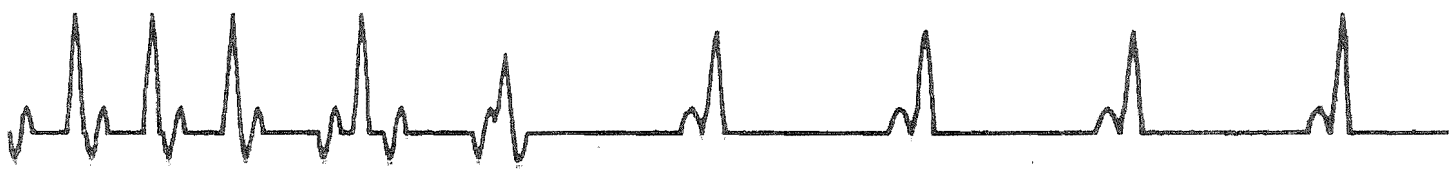
SI	N	SEC	TYPE
300	6	2	P
3200	1	2	P
270	6	2	P

SIMULATION TIME = 10 SEC



(a)





(b)

Fig. 5.12 Simulation for patient 2 (refer to text).
(a) First five seconds of simulation showing initiation of tachycardia by constant pacing at the atria.
(b) Following five seconds of simulation showing termination of tachycardia by overpacing at the atria.

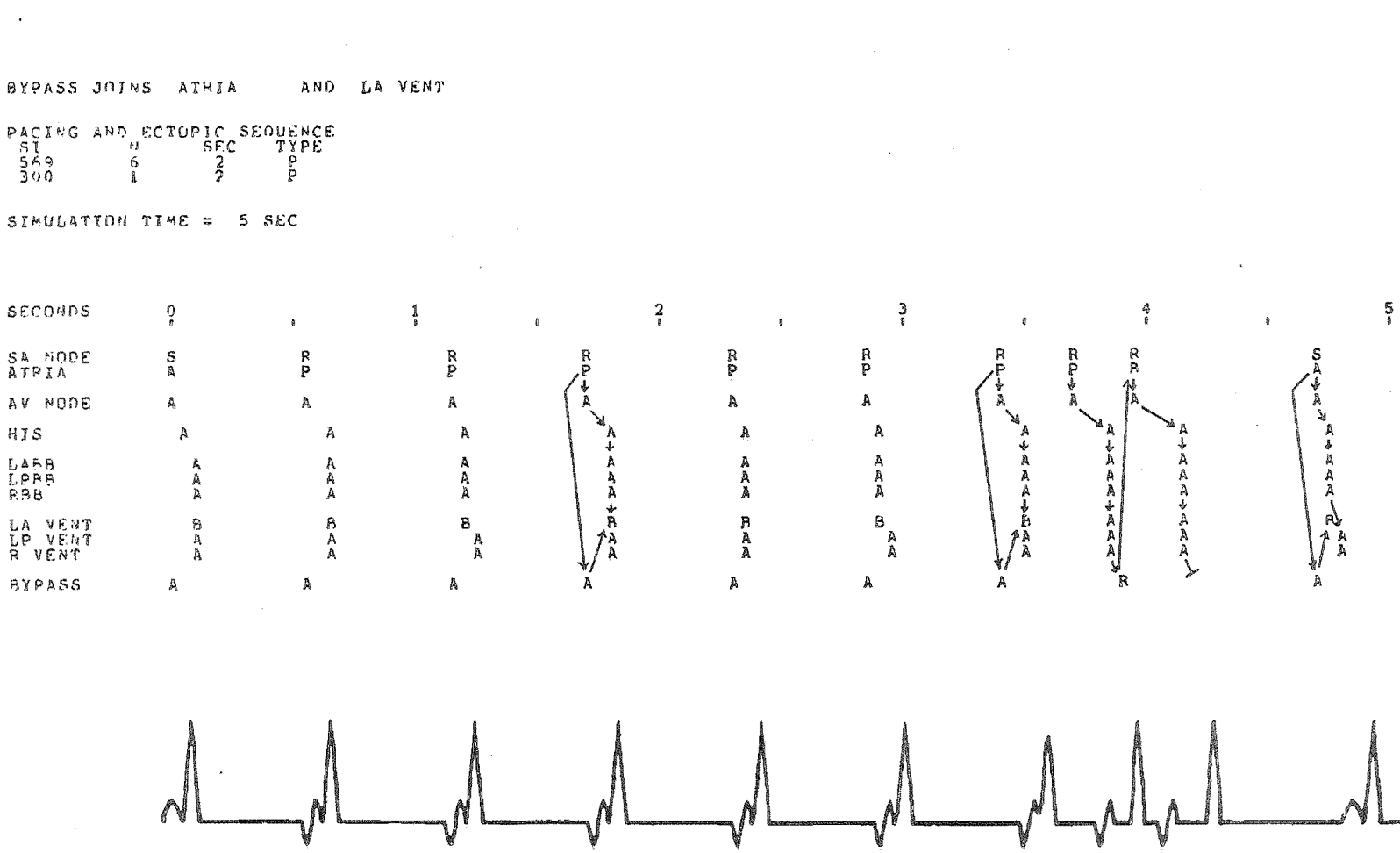


Fig. 5.13 Simulation of atrial extrastimulus testing for patient 2 (refer to text).

The response of the model to six stimuli of constant pacing at a CL of 300 ms is shown in Fig. 5.12(a). The ladder diagram in Fig. 5.12(a) shows that the pacing triggers a tachycardia, with a period of 298 ms, of the same type as that for patient 1 (Fig. 5.11). However this tachycardia does not spontaneously terminate. Fig. 5.12(b) is a continuation of the simulation shown in Fig. 5.12(a) and shows the tachycardia being overpaced at the atria at a CL of 270 ms. Overpacing begins 5 seconds after the start of the simulation and consists of six stimuli. Inspection of the ladder diagram in Fig. 5.12(b) shows that the high rate of the overpacing makes the AVN refractory, breaking the tachycardia loop and the model returns to sinus rhythm. Fig. 5.13 shows a simulation of an extrastimulus test with CL = 569 ms and SI = 300 ms. The extrastimulus does not trigger the tachycardia. Reference to the ladder diagram in Fig. 5.13 shows that this is because the accessory pathway is refractory to the high tachycardia rate, due to the large CL of the preceding stimuli.

5.5.4 Comments on the Examples

The examples described above show that the CCS model can be used to accurately simulate patients' heart rhythms under various conditions. The two patients studied both exhibit the same type of arrhythmia, however their arrhythmias are different with respect to their stability and the conditions under which they are precipitated. The simulations suggest possible electrophysiological bases for these differences.

Define a "tachycardia margin", m(i), for section i by

$$m(i) = \tau - e_i(\tau) \tag{5.17}$$

where τ is the tachycardia period and e_i is the ERP of section i. Clearly a stable tachycardia is only possible if $m(i) > 0$ for all tissue sections i in the re-entrant loop. In fact M, given by

$$M = \min_i [m(i)] \quad , \tag{5.18}$$

is a measure of a patient's "susceptibility to a tachycardia" - i.e. the patient is more susceptible for larger values of M. Inspection of tables 5.5 and 5.7, and making use of (5.2) and (5.17) gives

$$\left.\begin{aligned} m(AVN) &= -22 \text{ ms} \quad \text{for patient 1} \\ &= 28 \text{ ms} \quad \text{for patient 2} \end{aligned}\right\} \tag{5.19}$$

and

$$\left.\begin{aligned} m(\text{accessory pathway}) &= 36 \text{ ms} \quad \text{for patient 1} \\ &= 30 \text{ ms} \quad \text{for patient 2} \end{aligned}\right\} \quad (5.20)$$

Equations (5.19) and (5.20) show that patient 1 cannot support a stable tachycardia (because of the refractory properties of the AVN) whereas patient 2 can. It is for the same reason that patient 1's tachycardia is triggered by constant pacing whereas patient 2's is not. It is worth noting that the (clinically significant) result expressed by (5.19) is obtained from only a knowledge of the tachycardia period and the AVN pacing and extrastimulus measurements, by making use of the EP data processing described in §5.4.

Referring to Fig. 5.11, a transient tachycardia is precipitated in patient 1 by an extrastimulus for the following reason. The long CL preceding the extrastimulus decreases the ERP of the AVN so that it is, initially, able to conduct the fast tachycardia rate. However, after three beats at the fast rate, the AVN has adapted (because of the cumulative effect of CL) to the high rate and become refractory, terminating the tachycardia. The reason that an extrastimulus does not precipitate a tachycardia in patient 2 is (referring to Fig. 5.13) as follows. The long CL preceding the extrastimulus makes the accessory pathway sufficiently refractory to block the extrastimulus, preventing initiation of the tachycardia. Hence the difference in the response of the two patients to extrastimuli at their tachycardia periods is due to the different way in which refractoriness changes with CL for the AVN (i.e. $\beta < 0$) and the accessory pathway (i.e. $\beta > 0$).

5.6 DISCUSSION

The CCS model described in this chapter is based on parameters measured during EP studies rather than on the more fundamental electrophysiological properties used in other models (cf. references in §5.1). This allows a reasonably simple but realistic model of the whole CCS to be constructed which can be related to patients' EP data and used to simulate a wide range of arrhythmias. This approach to arrhythmia analysis seems to be new. The model has been implemented in such a way that it is easy for a cardiologist to use with very little training. The output display (particularly the ladder diagram) provides, in a familiar format,

detailed information which allows the simulation to be readily assessed. The examples presented in §5.5 show that the model can be used to assist the cardiologist to explore feasible mechanisms consistent with a patient's measured data. The model could also be used to predict the response of a patient to treatment with drugs or pacemakers.

One advantage of models of this type is that they may suggest (a) new ways of processing data to provide more useful information, and (b) new measurements which may be useful. For example the method described in §5.4 for determining β for the AVN from standard pacing and extrastimulus data allows the ERP of the AVN to be estimated under different stimulation conditions. The discussion in §5.4 suggests that useful information (for example, the parameters γ and b) on AVN conduction may be obtained by making accurate measurements during stable second degree block at various conduction ratios. This information is potentially very useful as it may allow one to predict the ability of the AVN to conduct high (and hence dangerous) rates to the ventricles in response to atrial fibrillation and other tachycardias. Further analysis of experimental data is required to assess the accuracy with which the CCS (particularly the AVN) has been modelled. Studies to determine if any of the model parameters described in §5.2 are related to particular forms of conduction system disease would also be useful.

6. CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

6.1 THE STRUCTURE OF DNA

X-ray diffraction is, at present, the only direct means available of imaging molecules at atomic resolution. However, as the work reported here has shown, the poor quality of DNA crystals (and the resulting low quality X-ray data) so far obtained precludes a unique structure determination. It is shown here that the gross features of the side-by-side model are consistent with the X-ray data (for the B form of DNA at least) and in fact agree with it slightly better than does the double helix model. It should be emphasised that structural refinements, based on the R-factor of an initial model do not provide a unique structure solution. However, with such low quality data, little else can be done. Refinement based on individual structure intensities, rather than the R-factor, may be preferable and should be explored.

The successful stereochemical refinement described here shows that the side-by-side model is a stereochemically viable structure. The use of theodolites to measure atomic coordinates of an initial physical model appears to be new and is a convenient means of obtaining such information from models of macromolecules.

The relevance of structure determinations of crystalline, i.e. *in vitro*, specimens of DNA to the *in vivo* structure is an important question. Structure determinations of short lengths of DNA are not necessarily representative of the *in vivo* structure. These difficulties could be alleviated if structure determinations of good crystals of DNA attached to protein, or of nucleosomes, could be made.

As described in §2.6, experiments with circular DNA provide information on the topological relationship between the two strands of DNA. Whether or not experiments based on this indirect technique can provide definitive structural information remains to be seen.

6.2 MACROSCOPIC INVERSE SCATTERING

The work reported here shows that the inverse scattering problem for plane stratified regions of variable refractive index is essentially

solved. However, little has been reported on the application of such techniques to measured scattering data. The results presented here of processing experimental data indicate the sensitivity of the algorithm to noisy data. However this algorithm appears to be more stable than many other reported techniques. The data pre-processing described here assists in stabilising the algorithm. Other methods of stabilising it by making use of any *a priori* information or measurement of the transmitted field should be explored.

Current methods of usefully imaging regions of arbitrary refractive index variation in more than one dimension apply only to weak inhomogeneities with radiation in either the low frequency (Born or Rayleigh-Gans approximation) or high frequency ("rays") regimes. Time domain methods which take account of multiple scattering are required. Similarly, attempts should be made to devise frequency domain methods of higher order, and which apply under wider conditions than, the Born approximation. While a general purpose technique would provide considerable insight, solutions that apply under different sets of restrictive conditions are likely to provide more tractable and useful solution techniques. Because the Born approximation applies to the Schrodinger equation under wider conditions than it does to the Helmholtz equation, use of the Chandrasakher transform and the Born approximation to process scattering data from regions of variable refractive index may provide improved reconstructions over those obtained by direct use of the Born approximation (refer to §1.5 and §1.9.2).

The effect of the nasal cavity on acoustical determination of vocal tract area functions should be investigated because it might provide useful information for speech analysis and synthesis. The inverse method for branched ducts described in §3.5 could provide a useful starting point for such an investigation.

Higher order solutions to the inverse eigenvalue problem require further investigation. The inverse eigenvalue problem for a circularly symmetric sphere is of vital importance in interpreting natural frequencies of oscillation of the earth. However the effect of having only one spectrum (i.e. for a single boundary condition at the surface - refer to §1.10) available for measurement, on uniqueness of the solution, is not well understood and should be further investigated.

6.3 CARDIAC ELECTROPHYSIOLOGICAL MODELLING

The model of the cardiac conduction system described here is potentially useful in a clinical setting because it is based on physiological parameters which are routinely measured. This approach to modelling cardiac rhythms does not yet appear to have been explored. However, because of the complexity of cardiac conduction and the small amount of data obtained during routine clinical investigations, it is unlikely that modelling can be used to supply specific diagnostic information on arrhythmias. The model described here is, however, potentially useful in suggesting possible arrhythmia mechanisms to the cardiologist. Trials in which a cardiologist uses the model are necessary to assess this potential.

In order to assess, and possibly improve, the accuracy of the model, more definitive experimental data on the dependence of the atrio-ventricular nodal refractory curve on cycle length are required. The model suggests measurements (made during second degree block) which may assist in characterising the conduction properties of the atrio-ventricular node. An experimental study is needed to determine the usefulness of these measurements. A clinical study is required to assess the usefulness of the tachycardia index, M, to measure patients' susceptibility to tachycardias.

One of the most promising potential applications of the model is as a teaching aid. With the addition of suitable interactive software, the model could be used as the basis of a computer aided learning program. An interaction with the program could involve simulation of the effects of different drugs and diseases, and also nervous and hormone effects. For such a program it would be essential to include the autonomic effect which couples sinus node rate to atrio-ventricular nodal refractoriness (mentioned in §5.2) in the model.

APPENDIX 1

DECONVOLUTION METHODS

Let f(x) be the convolution of g(x) with the point spread function h(x) so that

$$f(x) = g(x) \otimes h(x) + n(x)$$
$$= \int_{-\infty}^{\infty} g(y)\, h(x-y)\, dy + n(x) \qquad (A1.1)$$

where n(x) represents the measurement noise. The functions f(x) and g(x) are referred to as the image and the object respectively in image processing terminology. Deconvolution is the process of recovering g(x) when f(x) and h(x) are known. The bases of the two deconvolution techniques invoked in §3.3, which may be called subtractive and multiplicative methods, are outlined here.

Subtractive Deconvolution

The method clean (Schwarz, 1978; Bates *et al.*, 1982a) is a subtractive deconvolution technique used originally for processing radiointerferometery images. It is most useful for objects which consist of isolated peaks. The strategy of the procedure is to subtract scaled versions of the point spread function from the image until only the measurement noise remains. In order to simplify the subsequent discussion, h(x) is shifted and scaled so that

$$\max_x |h(x)| = h(0) = 1 \; . \qquad (A1.2)$$

The residual image, $f_n(x)$, after the n^{th} iteration is given by

$$f_n(x) = f_{n-1}(x) - \alpha\, f_{n-1}(x_{max})\, h(x - x_{max}) \; , \qquad (A1.3)$$

where α is a constant called the loop gain and x_{max} is given by

$$\max_x |f_{n-1}(x)| = f_{n-1}(x_{max}) \; . \qquad (A1.4)$$

The estimated object, $g_n(x)$, after the n^{th} iteration is given by

$$g_n(x) = g_{n-1}(x) + \alpha\, f_{n-1}(x_{max})\, \delta(x - x_{max}) \quad . \tag{A1.5}$$

The procedure is stopped when

$$(\max_x |f_n(x)|)/(\max_x |f(x)|) < \gamma \tag{A1.6}$$

where γ is the estimated noise level. Fig. A1.1 shows f(x), h(x) and the estimated g(x) for the example in Fig. 3.3.

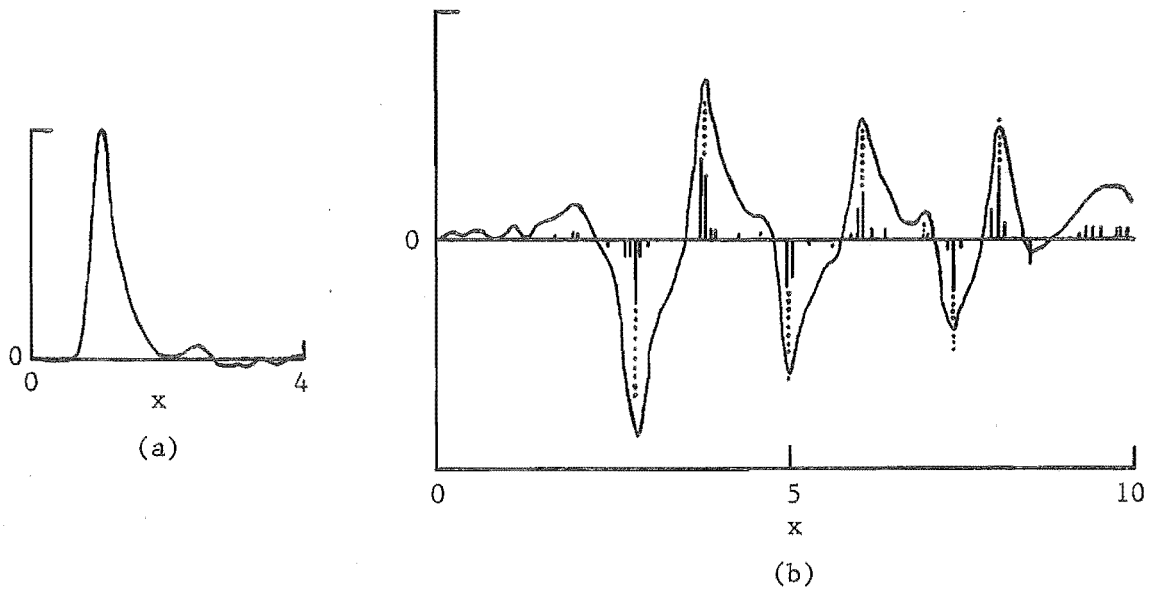


Fig. A1.1 "Cleaning" the reflected field for the example in Fig. 3.3.
(a) h(x) (the incident field).
(b) g(x) (vertical dotted lines) (the calculated impulse response).
f(x) (solid curve) (the measured reflected field).
$g_N(x)$ (vertical solid lines) (the estimated impulse response for $\alpha = 0.08$, $\gamma = 0.08$, $N = 270$).

Multiplicative Deconvolution

Making use of the convolution theorem (Bracewell, 1978, ch. 6) allows (A1.1) to be written as

$$F(u) = G(u)H(u) + N(u) \tag{A1.7}$$

where upper case letters denote the FT and u is the transform variable. Dividing (A1.7) by H(u) gives

$$F(u)/H(u) = G(u) + N(u)/H(u) \quad . \tag{A1.8}$$

The difficulty with using (A1.8) to estimate G(u) is that the noise is amplified when H(u) is small. It can be shown (cf. Helstrom, 1967) that the best least squares estimate, $\tilde{G}(u)$, of G(u) is given by

$$\tilde{G}(u) = F(u)W(u) \tag{A1.9}$$

where W(u), called the Wiener filter, is given by

$$W(u) = H^*(u)\, P_f(u)/[|H(u)|^2\, P_f(u) + P_n(u)] \tag{A1.10}$$

and $P_f(u)$ and $P_n(u)$ are the power spectra of the image and the noise respectively. Usually, in practice, the dependence of $P_f(u)$ and $P_n(u)$ on u is unknown so that $P_n(u)/P_f(u)$ is often replaced by a constant so that

$$W(u) = H^*(u)/(|H(u)|^2 + \phi) \tag{A1.11}$$

where ϕ is an estimate of the square of the noise to signal ratio. Hence g(x) is estimated by using (A1.9) and (A1.11) to determine $\tilde{G}(u)$ and taking the inverse FT. Fig. A1.2 shows f(x), h(x) and g(x) and an estimate, $\tilde{g}(x)$, of g(x) obtained by Wiener filtering the reflected field calculated for the profile shown in Fig. 3.2(b).

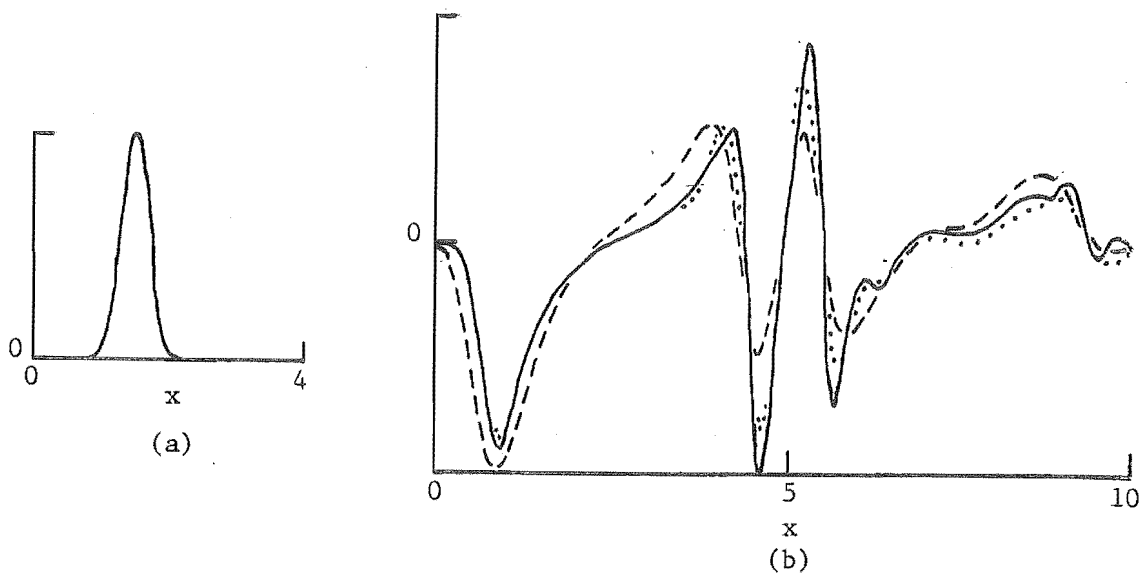


Fig. A1.2 Wiener filtering the reflected field calculated for the profile shown in Fig. 3.2(b) for $\gamma = 0.1$ and $L = 7$ (see §3.4).
(a) h(x) (the incident field).
(b) g(x) (——) (the actual impulse response)
f(x) (– – –) (the reflected field)
$\tilde{g}(x)$ (◦◦◦) (the estimated impulse response for $\phi = 0.01$).

APPENDIX 2

TIME DOMAIN REFLECTOMETER EXPERIMENTS

The time domain reflectometer (TDR) (Hewlett Packard models HP4151A and HP140B) and associated equipment used to obtain the experimental data described in §3.4 and §3.5 are shown in Fig. A2.1. A variable length short-circuited T-section was used to produce a pulse from the voltage step provided by the TDR. The pulse was launched in the coaxial test line and the reflected waveform captured by the TDR and subsequently digitized and stored in the computer.

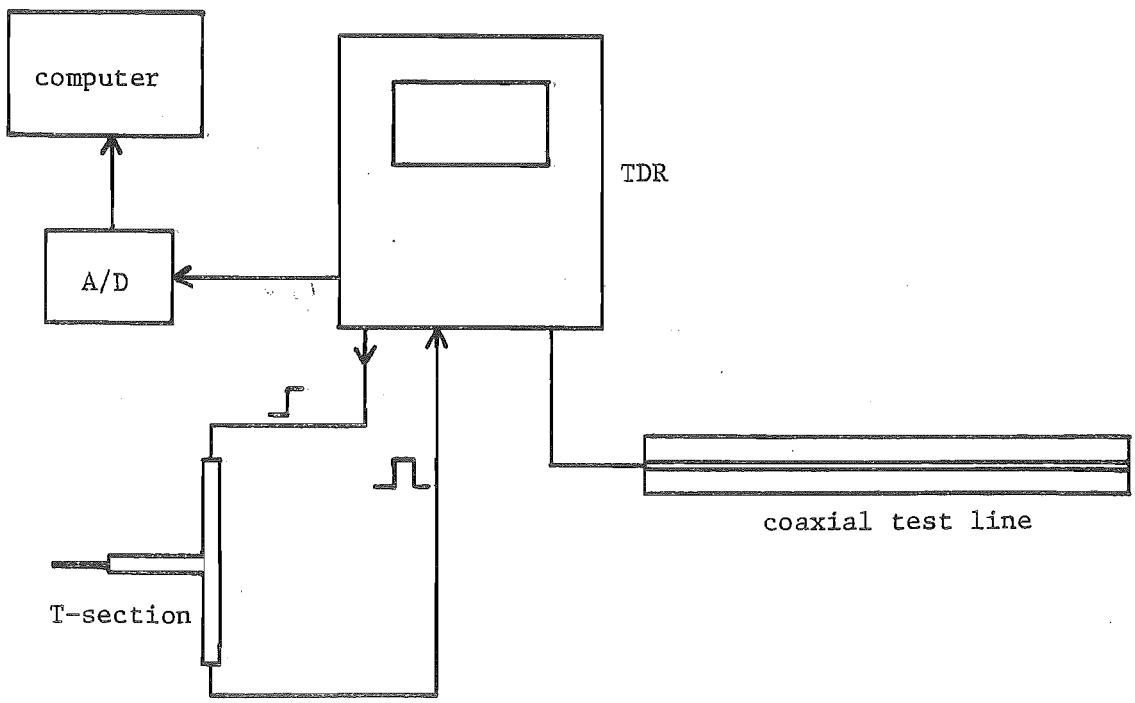


Fig. A2.1 Apparatus for TDR experiments.

It was found necessary to use sufficiently long lengths of coaxial cable feeding the T-section that transients reflected by the connectors did not interfere with the wanted signal. The cross-sectional dimensions of the line were sufficiently small that only the TEM mode (see §1.4.1) propagates. Perspex beads were inserted in the line to provide a variable refractive index. The incident pulse at the input to the line was measured by placing a short circuit there and capturing the reflected pulse. The incident pulse had a duration of approximately 0.8 ns which corresponds to a length of 16 cm (between half power points) in the air filled line. The actual experimental arrangement used for the measurements reported in §3.4 and §3.5 was originally set up by Coronno (1980) and McNeill (1980).

APPENDIX 3

BRANCHED NETWORKS

Branched networks considered here consist of interconnected sections of ducts or transmission lines. The impedance of each section is constant. To simplify the discussion of the inversion procedure described in §3.5, a particular path through the network is chosen as the main line as shown in Fig. A3.1. Position along the main line is identified by the coordinate x. Measurements are made at the observation point x = 0. The structure of the network refers to the branching pattern of each branch.

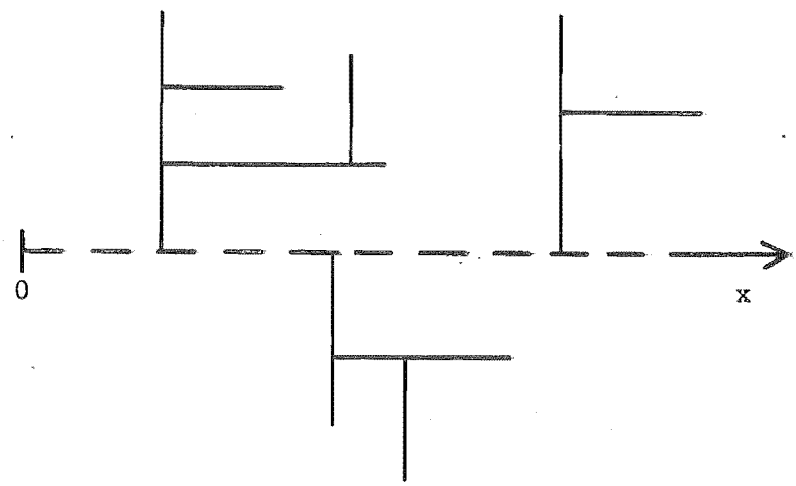


Fig. A3.1 A general branched network consisting of a main line (---) and branches (——).

Consider a branch of impedance ζ on a line of impedance ζ_o. The branch may be either in parallel or in series with the main line as shown in Fig. A3.2. Requiring that the voltage and current are continuous at the junctions implies that the voltage reflection coefficients for waves on the line are given by

$$\Gamma_p = -\zeta_o/(\zeta_o + 2\zeta) \tag{A3.1}$$

and

$$\Gamma_s = \zeta/(2\zeta_o + \zeta) \tag{A3.2}$$

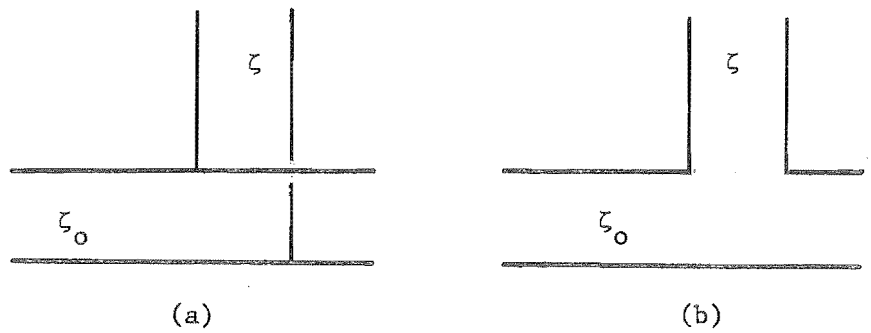


Fig. A3.2 (a) Parallel and (b) series branches.

for parallel and series branches respectively. Inspection of (A3.1) and (A3.2) shows that

$$\Gamma_p < 0 \quad \text{and} \quad \Gamma_s > 0 \tag{A3.3}$$

so that, given ζ_o, the impedance of the branch, and whether it is in parallel or series with the line, can be calculated from the reflection coefficient.

A branch of infinite length (or terminated in a matched load) has the same effect as a localised loss. Hence localised losses can be treated in the same manner as described above. A localised loss may be either in parallel or series with the line. For an EM transmission line, parallel and series loss could be due to non-zero conductivity of the medium in which the line is embedded and non-zero conductor resistance respectively. For an acoustic duct, parallel and series loss could be due to a hole in the duct wall and viscous energy loss due to a sudden change in duct cross-section respectively (Morse and Ingard, 1968, §9.1).

APPENDIX 4

A POWER SERIES FORMULATION OF THE INVERSE EIGENVALUE PROBLEM

Consider the inverse eigenvalue problem (as described in §1.10) for the HE

$$\partial^2\psi/\partial x^2 + k^2\mu\psi = 0 \tag{A4.1}$$

on the interval (0,1) where $\mu = \mu(x)$ is the square of the refractive index profile. The wavefunction satisfies the boundary condition

$$\partial\psi(0,k)/\partial x = 0 \tag{A4.2}$$

at $x = 0$ and measurements are made at $x = 1$. The profile μ is expanded as a power series,

$$\mu(x) = \sum_{n=o}^{\infty} \mu_n x^{2n} \quad , \quad 0 \leq x \leq 1 \ . \tag{A4.3}$$

The wavefunction is expanded as a power series in k,

$$\psi(x,k) = 1 + \sum_{\ell=1}^{\infty} B_\ell(x)\, k^{2\ell} \ , \tag{A4.4}$$

and substituting (A4.4) into (A4.1) gives

$$\left.\begin{aligned} B_1''(x) &= -\mu(x) \\ \text{and} \\ B_{\ell+1}''(x) &= -\mu(x)\, B_\ell(x) \end{aligned}\right\} \tag{A4.5}$$

where a prime denotes the derivative with respect to x. Inspection of (A4.4) and (A4.5) shows that $B_\ell(x)$ is of the form

$$B_\ell(x) = x^{2\ell} \sum_{m=o}^{\infty} b_{\ell m} x^{2m} \quad , \quad 0 \leq x \leq 1 \tag{A4.6}$$

where the $b_{\ell m}$ are constants. Substituting from (A4.3) and (A4.6) into (A4.5) and equating like powers of x gives

$$2(\ell+m+1)(2\ell+2m+1)\, b_{\ell+1,m} = \sum_{n=o}^{m} b_{\ell,m-n}\, \mu_n \ . \tag{A4.7}$$

Inspection of (A4.4) shows that the wavefunction at the measurement point $x = 1$ is given by

$$\psi(1,k) = 1 + \sum_{\ell=1}^{\infty} B_\ell \, k^{2\ell}, \tag{A4.8}$$

where $B_\ell = B_\ell(1)$, which can be written as

$$\psi(1,k) = \prod_{\ell=1}^{\infty} (1 - k^2/k_\ell^2) \ . \tag{A4.9}$$

The k_ℓ are the measured short circuit resonant frequencies and so, using (A4.8) and (A4.9), the constants B_ℓ can be computed from the measurements. Similar reasoning shows that the constants $B'_\ell = B'_\ell(1)$ can be computed from the open circuit resonances. Making use of (A4.6) shows that

$$\left.\begin{aligned} B_\ell &= \sum_{m=o}^{\infty} b_{\ell m} \\ \text{and} \qquad B'_\ell &= \sum_{m=o}^{\infty} 2(m+\ell) \, b_{\ell m} \ . \end{aligned}\right\} \tag{A4.10}$$

If L both short circuit and open circuit resonances are measured and the summations in (A4.10) are truncated to m = L, then (A4.10) provides 2L linear equations in the $b_{\ell m}$. In addition, (A4.7), for $0 \leq \ell \leq L-1$ and $0 \leq m \leq L$, provides $L(L+1)$ quadratic equations in the $b_{\ell m}$ and μ_n. In total there are $L(L+3)$ equations in $(L+1)^2$ unknowns which include μ_n for $0 \leq n \leq L$. Solution of this (overdetermined) set of equations for the μ_n, in principle, solves the inverse eigenvalue problem. However, because some of the equations are quadratic, the solution is not straightforward and may not be unique. Also, power series expansion of the wavefunction is not particularly efficient since oscillatory functions require a large number of terms for reasonable accuracy.

REFERENCES

Akhtar M., Damato A.N., Batsford W.P., Ruskin J.N. and Ogunkelu J.B. (1975) A Comparative Analysis of Antegrade and Retrograde Conduction Patterns in Man, Circ., 52, 766-778.

Aki K. and Richards P.G. (1980) Quantitative Seismoloty - Theory and Methods, vols 1 and 2, Freeman and Co., San Francisco.

Arnott S. (1970) The Geometry of Nucleic Acids, in Butler J.A.V. and Noble D. (eds), Prog. in Biophys. and Molec. Biol., 21, 265-319.

Arnott S. and Chandrasekaran R. (1981) Fibrous Polynucleotide Duplexes Have Very Polymorphic Secondary Structures, in Sarma R.H. (ed.), Biomolecular Stereodynamics (Proc. Second SUNYA Conversation in the Discipline Molecular Stereodynamics), I(II), 99-122.

Arnott S. and Hukins D.W.L. (1972a) Optimised Parameters for A-DNA and B-DNA, Biochem. Biophys. Res. Comm., 47, 1504-1509.

Arnott S. and Hukins D.W.L. (1972b) The Dimensions and Shapes of the Furanose Rings in Nucleic Acids, Biochem. J., 130, 453-465.

Arnott S. and Hukins D.W.L. (1973) Refinement of the Structure of B-DNA and Implications for the Analysis of X-ray Diffraction Data from Fibres of Biopolymers, J. Mol. Biol., 81, 93-105.

Arnott S., Chandrasekaran R., Birdsall D.L., Leslie A.G.W. and Ratliff R.L. (1980) Left-handed DNA Helices, Nature, 283, 743-745.

Backus G. and Gilbert F. (1968) The Resolving Power of Gross Earth Data, Geophys. J. Roy. Astr. Soc., 16, 169-205.

Baltes H.P. (ed.) (1978) Inverse Source Problems in Optics, Topics in Current Physics, vol. 9, Springer-Verlag, N.Y.

Baltes H.P. (ed.) (1980) Inverse Scattering Problems in Optics, Topics in Current Physics, vol. 20, Springer-Verlag, N.Y.

Barcilon V. (1975) Well-posed Inverse Eigenvalue Problems, Geophys. J. Roy. Astr. Soc., 42, 375-383.

Barcilon V. (1978) On the Information Content of Natural Frequency Spectra Associated with Different Angular Numbers, Geophys. J. Roy. Astr. Soc., 53, 623-641.

Barcilon V. (1979) Ideal Solution of an Inverse Normal Mode Problem with Finite Spectral Data, Geophys. J. Roy. Astr. Soc., 56, 399-408.

Bates J.H.T., McKinnon A.E. and Bates R.H.T. (1982a) Subtractive Image Restoration I: Basic Theory, Optik, in press.

Bates J.H.T., McKinnon A.E. and Bates R.H.T. (1982b) Subtractive Image Restoration II: Comparison with Multiplicative Deconvolution, Optik, in press.

Bates J.H.T., Fright W.R., Millane R.P., Seagar A.D., McKinnon A.E. and Bates R.H.T. (1982c) Subtractive Image Restoration III: Some Practical Applications, to be submitted to Optik.

Bates R.H.T. (1975) Global Solution to the Scalar Inverse Scattering Problem, J. Phys. A: Math. Gen., 8, L80-L82.

Bates R.H.T. (1977) Towards Useful Exact Remote Sensing Theory, Institute of Theoretical Physics, Chalmers Tekniska Högskolan, Göteborg, Sweden, Report no. 77-16.

Bates R.H.T. (1978) On Phase Problems: I and II, Optik, 51, 161-170 and 223-234.

Bates R.H.T. and Millane R.P. (1981) Time Domain Approach to Inverse Scattering, IEEE Trans. Ant. Prop., AP-29, 359-363.

Bates R.H.T. and Ng F.L. (1972) Polarization-source Formulation of Electromagnetism and Dielectric-loaded Waveguides, Proc. IEE, 119, 1568-1574.

Bates R.H.T. and Wall D.J.N. (1976) Chandrasekhar Transformations Improve Convergence of Computations of Scattering from Linearly Stratified Media, IEEE Trans. Ant. Prop., AP-24, 251-253.

Bates R.H.T. and Wall D.J.N. (1977) Null Field Approach to Scalar Diffraction, Phil. Trans. Roy. Soc. London, 287, 45-114.

Bates R.H.T., Boerner W.M. and Dunlop G.R. (1976) An Extended Rytov Approximation and its Significance for Remote Sensing and Inverse Scattering, Optics Comm., 18, 421-423.

Bates R.H.T., Lewitt R.M., Rowe C.H., Day J.P. and Rodley G.A. (1977) On the Structure of DNA, J. Roy. Soc. N.Z., 7, 273-301.

Bates R.H.T., McKinnon G.C. and Millane R.P. (1978) A New Look at B-DNA Diffraction Data, Research Report, Elec. Eng. Dept., University of Canterbury, New Zealand.

Bates R.H.T., McKinnon G.C., Millane R.P. and Rodley G.A. (1980) Revised Interpretations of the Available X-ray Data for B-DNA, Pramāna, 14, 233-252.

Bennett C.L. (1981) Time Domain Inverse Scattering, IEEE Trans. Ant. Prop., AP-29, 213-219.

Berbari E.J., Lazzara R. and Scherlag B.J. (1979) The Effects of Filtering the His-Purkinje System Electrocardiogram, IEEE Trans. Biomed. Eng., BME-26, 82-85.

Bhereur P., Roberge F.A. and Nadeau R.A. (1971) A Simulation Unit for Cardiac Arrhythmias, Med. Biol. Eng., 9, 13-21.

Boerner W.M., Jordan A.K. and Kay I.W. (eds.) (1981) IEEE Trans. Ant. Prop., Special Issue on Inverse Methods in Electromagnetics, AP-29, 185-417.

Bolomey J.C., Lesselier D., Pichot C. and Tabbara W. (1981) Spectral and Time Domain Approaches to Some Inverse Scattering Problems, IEEE Trans. Ant. Prop., AP-29, 206-212.

Bones P.J. (1981) Contributions to Electrocardiographic Science, PhD thesis, University of Canterbury, New Zealand.

Borg G. (1946) Eine Umkehrung der Sturm-Liouvilleschen Eigenwertaufgabe, Acta Math. 78, 1-96 (in German), English abstract in Math. Rev. (1946), 7, 382.

Bracewell R.N. (1978) The Fourier Transform and its Applications (2nd ed.), McGraw-Hill, N.Y.

Bram S. (1977) Private Communication.

Bremmer H. (1951) The W.K.B. Approximation as the First Term of a Geometric-Optical Series, Comm. Pure App. Math., 4, 105-115.

Bullen K.E. (1963) An Introduction to the Theory of Seismology, Cambridge University Press.

Chadan K. and Sabatier P.C. (1977) Inverse Problems in Quantum Scattering Theory, Springer-Verlag, N.Y.

Chernov L. (1967) Wave Propagation in a Random Medium, Dover, N.Y.

Colin L. (ed.) (1972) Mathematics of Profile Inversion, NASA Technical Memorandum, TM X-62, 150.

Coronno A.C. (1980) Remote Probing Experiment - Time Domain Analysis, Elec. Eng. Dept. Project Report, University of Canterbury, New Zealand.

Cowley J.M. (1975) Diffraction Physics, North-Holland, Amsterdam.

Crick F.H.C. and Klug A. (1975) Kinky Helix, Nature, 255, 530-533.

Crick F.H.C., Wang J.C. and Bauer W.R. (1979) Is DNA Really a Double Helix?, J. Mol. Biol., 129, 449-461.

Curry P.V.L. (1975) Fundamentals of Arrhythmias: Modern Methods of Investigation, in Krikler and Goodwin (1975), 39-80.

Denes P., Wu D., Dhingra R., Peitras R.J. and Rosen K.M. (1974) The Effects of Cycle Length on Cardiac Refractory Periods in Man, Circulation, 49, 32-41.

Deschamps G.A. and Cabayan H.S. (1972) Antenna Synthesis and Solution of Inverse Problems by Regularization Methods, IEEE Trans. Ant. Prop., AP-20, 268-274.

Donohue J. (1969) Fourier Analysis and the Structure of DNA, Science, 165, 1091-1096.

Dreifus L.S., Watanabe Y., Dreifus H.N. and Azevedo I.D. (1976) The Effect of Antiarrhythmic Agents on Impulse Formation and Impulse Conduction, in Wellens H.J.J. (ed.), The Conduction System of the Heart, Lea and Febiger, Netherlands.

Ferrier G.R. and Dresel P.E. (1973) Role of the Atrium in Determining the Functional and Effective Refractory Periods and the Conductivity of the Atrioventricular Transmission System, Circ. Res., 33, 375-385.

Feughelman M., Langridge R., Seeds W.E., Stokes A.R., Wilson H.R., Hooper C.W., Wilkins M.H.F., Barclay R.K. and Hamilton L.D. (1955) Molecular Structure of Deoxyribose Nucleic Acid and Nucleoprotein, Nature, 175, 834-838.

Fienup J.R. (1978) Reconstruction of an Object from the Modulus of its Fourier Transform, Opt. Lett., 3, 27-29.

Fisher J.D. (1981) Role of Electrophysiologic Testing in the Diagnosis and Treatment of Patients with Known and Suspected Bradycardias and Tachycardias, Prog. Cardiovasc. Diseases, 24, 25-90.

Flanigan L.K. and Swain H.H. (1967) Computer Simulation of A-V Nodal Conduction, Uni. Michigan Med. Centre J., 33, 234-241.

Gallagher J.J., Pritchett E.L.C., Sealy W.C., Kasell J. and Wallace A.G. (1978) The Preexcitation Syndromes, Prog. Cardiovasc. Diseases, 20, 285-327.

Gel'fand I.M. and Levitan B.M. (1951) On the Determination of a Differential Equation by its Spectral Function, Izv. Akad. Nauk. SSSR Ser. Mat., 15, 309-360 (in Russian), English translation in Am. Math. Soc. Transl. (1955), 1, 253-304.

Gerver M.L. (1970) Inverse Problem for the One-dimensional Wave Equation, Geophys. J. Roy. Astr. Soc., 21, 337-357.

Gilbert F. and Dziewonski A.M. (1975), An Application of Normal Mode Theory to the Retrieval of Structural Parameters and Source Mechanisms from Seismic Spectra, Phil. Trans. Roy. Soc. London, 278, 187-269.

Goldman M.J. (1979) Principles of Clinical Electrocardiography (10th ed.), Lange Medical Publications, Calif.

Gopinath B. and Sondhi M.M. (1970) Determination of the Shape of the Human Vocal Tract from Acoustical Measurements, Bell Sys. Tech. J., 49, 1195-1214.

Hardcastle W.J. (1976) Physiology of Speech Production - An Introduction for Speech Scientists, Academic Press, N.Y.

Haschemeyer A.E.V. and Rich A. (1967) Nucleoside Conformations: an Analysis of Steric Barriers to Rotation about the Glycosidic Bond, J. Mol. Biol., 27, 369-384.

Heading J. (1962) An Introduction to Phase-Integral Methods, Methuen, London.

Heethaar R.M., Denier van der Gon J.J. and Meijler F.L. (1973a) Mathematical Model of A-V Conduction in the Rat Heart, Cardiovas. Res., 7, 105-114.

Heethaar R.M., de vos Burchart R.M., Denier van der Gon J.J. and Meijler F.L. (1973b) A Mathematical Model of A-V Conduction in the Rat Heart. II. Quantification of Concealed Conduction, Cardiovas. Res., 7, 542-556.

Heethaar R.M., Denier van der Gon J.J. and Meijler F.L. (1973c) Interpretation of Some Properties of A-V Conduction with the help of Analog Simulation, Europ. J. Card., 1/1, 87-93.

Helstrom C.W. (1967) Image Restoration by the Method of Least Squares, J. Opt. Soc. Am., 57, 297-303.

Hingerty B. (1979) Geometry of Experimentally Observed RNA Residues in tRNA and Dinucleoside Monophosphates: The Effect of Small Variations in the Backbone Angles, Biopolymers, 18, 1901-1915.

Hodgkin A.L. and Huxley A.F. (1952) A Quantitative Description of Membrane Current and its Application to Conduction and Excitation in Nerve, J. Physiol. (London), 117, 500-544.

Janse M.J., van der Steen A.B.M., van Dam R.T. and Durrer D. (1969) Refractory Period of the Dog's Ventricular Myocardium Following Sudden Changes in Frequency, Circ. Res., 24, 251-262.

Jones D.S. (1964) The Theory of Electromagnetism, Pergamon Press, Oxford.

Jordan A.K. and Ahn S. (1979) Inverse Scattering Theory and Profile Inversion, Proc. IEE, 126, 945-950.

Jordan E.C. and Balmain K.G. (1968) Electromagnetic Waves and Radiating Systems (2nd ed.), Prentice-Hall, Englewood Cliffs, N.J.

Joyner R.W., Ramon F. and Moore J.W. (1975) Simulation of Action Potential Propagation in an Inhomogeneous Sheet of Coupled Excitable Cells, Circ. Res., 36, 654-661.

Karbowiak A.E. (1957) Propagation of Transients in Waveguides, Proc. IEE, 104C, 339-348.

Kay I. (1960) The Inverse Scattering Problem when the Reflection Coefficient is a Rational Function, Comm. Pure Appl. Math., 13, 371-398.

Kay I. (1972) The Inverse Scattering Problem for Transmission Lines, in Colin (1972), ch. 6.

Kay I. and Moses H.E. (1961) A Simple Verification of the Gelfand-Levitan Equation for the Three-Dimensional Scattering Problem, Comm. Pure and Appl. Math., 14, 435-445.

Keller J.B. (1969) Accuracy and Validity of the Born and Rytov Approximations, J. Opt. Soc. Am., 59, 1003-1004.

Krein M.G. (1952) On Inverse Problems for an Inhomogeneous Cord, Doklady Akad. Nauk. SSSR, 82, 669-672 (in Russian), English abstract in Math. Rev. (1953), 14, 649.

Krikler D.M. and Goodwin J.F. (eds.) (1975) Cardiac Arrhythmias: The Modern Electrophysiological Approach, Saunders, London.

Krikler D.M. and Wellens H.J.J. (1975) The Wolff-Parkinson-White and Related Syndromes, in Krikler and Goodwin (1975), 144-181.

Langendorf R. (1948) Concealed A-V Conduction: The Effect of Blocked Impulses on the Formation and Conduction of Subsequent Impulses, Am. Heart J., 35, 542-552.

Lesselier D. (1978) Determination of Index Profiles by Time Domain Reflectometry, J. Optics (Paris), 9, 349-358.

Levitan B.M. (1964) On the Determination of a Sturm-Liouville Equation by Two Spectra, Izv. Akad. Nauk. SSSR Ser. Mat., 28, 63-78 (in Russian), English translation in Am. Math. Soc. Transl. (1968), 68, 1-20.

MacGillavry C.H. and Bruins E.M. (1948) On the Patterson Transforms of Fibre Diagrams, Acta Cryst., 1, 156-158.

McKinnon G.C. (1980) Contributions to Imaging, PhD Thesis, University of Canterbury, New Zealand.

McNeill R.G. (1980) Remote Probing Experiment, Elec. Eng. Dept. Project Report, University of Canterbury, New Zealand.

Mermelstein P. (1967) Determination of the Vocal-Tract Shape from Measured Formant Frequencies, J. Acoust. Soc. Am., 41, 1283-1294.

Millane R.P. and Bates R.H.T. (1981) Inverse Methods for Branched Ducts and Transmission Lines, Proc. IEE on Communications, Radar and Signal Processing, in press.

Millane R.P. and Rodley G.A. (1981) Stereochemical Details of the Side-by-Side Model for DNA, Nucleic Acids Res., 9, 1765-1773.

Millane R.P., Bones P.J., Ikram H. and Bates R.H.T. (1980) A Computer Model of Cardiac Conduction, Australasian Phys. Eng. Sci. Med., 3, 205-209.

Millane R.P., Rodley G.A. and Rodley G.F. (1982) Refinement of the Side-by-Side Model for DNA, J. Roy. Soc. N.Z., in press.

Millar R.N. (1979) A Simple Model of Delay, Block and One Way Conduction in Purkinje Fibers, J. Math. Biol., 7, 385-398.

Mittra R., Schaubert D.H. and Mostafavi M. (1972) Some Methods for Determining the Profile Functions of Inhomogeneous Media, in Colin (1972), ch. 8.

Morse P.M. (1948) Vibration and Sound, McGraw-Hill, N.Y.

Morse P.M. and Feshbach H. (1953) Methods of Theoretical Physics, parts 1 and 2, McGraw-Hill, N.Y.

Morse P.M. and Ingard K.U. (1968) Theoretical Acoustics, McGraw-Hill, N.Y.

Newton R.G. (1966) Scattering Theory of Waves and Particles, McGraw-Hill, N.Y.

Newton R.G. (1974) The Gel'fand Levitan Method in the Inverse Scattering Problem, in Lavita J.A. and Marchand J.P. (eds.), Scattering Theory in Mathematical Physics, NATO Advanced Study Institutes Series C9, D. Reidel, Holland.

Noble D. (1962) A Modification of the Hodgkin-Huxley Equations Applicable to Purkinje Fibre Action and Pacemaker Potentials, J. Physiol., 160, 317-352.

Porgorzelski W. (1966) Integral Equations and their Applications, Pergamon Press, Oxford.

Ramachandran G.N. and Srinivasan R. (1970) Fourier Methods in Crystallography, Wiley, N.Y.

Rodgers C.D. (1976) Retrieval of Atmospheric Temperature and Composition from Remote Measurements of Thermal Radiation, Rev. Geophys. Space Phys., 14, 609-624.

Rodley G.A., Scobie R.S., Bates R.H.T. and Lewitt R.M. (1976) A Possible Conformation for Double Stranded Polynucleotides, Proc. Nat. Acad. Sci. USA, 73, 2959-2963.

Sabatier P.C. (1978) Inverse Scattering Problems for Nonlinear Applications, in Barut A.O. (ed.), Nonlinear Equations in Physics and Mathematics, NATO Advanced Study Institures Series, C40, D. Reidel, Holland.

Schamroth L. (1971) The Disorders of Cardiac Rhythm, Blackwell Sci. Pubs., Oxford.

Schroeder M.R. (1967) Determination of the Geometry of the Human Vocal Tract by Acoustic Measurements, J. Acoust. Soc. Am., 41, 1002-1010.

Schwarz U.G. (1978) Mathematical-Statistical Description of the Iterative Beam Removing Technique (method CLEAN), Astr. Astrophys., 65, 345-356.

Sherwood D. (1976) Crystals, X-rays and Proteins, Longman, London.

Simpson M.B., Spear J.F. and Moore E.N. (1978) Electrophysiologic Studies on Atrioventricular Nodal Wenckebach Cycles, Am. J. Cardiol., 41, 244-258.

Simpson M.B., Spear J. and Moore E.N. (1979) The Relationship between Atrioventricular Nodal Refractoriness and the Functional Refractory Period in the Dog, Circ. Res., 44, 121-126.

Spurrell R.A.J. (1975a) Surgery for Ventricular Tachycardia, in Krikler and Goodwin (1975), 195-207.

Spurrell R.A.J. (1975b) Artificial Cardiac Pacemakers, in Krikler and Goodwin (1975), 238-258.

Stansfield E.V. and Bogner R.E. (1973) Determination of Vocal-Tract-Area Function from Transfer Impedance, Proc. IEE, 120, 153-158.

Teague S., Collins S., Wu D., Denes P., Rosen K. and Arzbaecher R. (1976) A Quantitative Description of Normal AV Nodal Conduction Curve in Man, J. App. Physiol., 40, 74-78.

Tijhuis A.G. (1981) Iterative Determination of Permittivity and Conductivity Profiles of a Dielectric Slab in the Time Domain, IEEE Trans. Ant. Prop., AP-29, 239-245.

Van Capelle F.J.L. and Durrer D. (1980) Computer Simulation of Arrhythmias in a Network of Coupled Excitable Elements, Circ. Res., 47, 454-466.

Vezzetti D.J. and Aks S.O. (1979) Reconstructions from Scattering Data: Analysis and Improvements of the Inverse Born Approximation, Ultrasonic Imaging, 1, 333-345.

Wang A.H., Nathans J., van der March G., van Boom J.H. and Rich A. (1979) Molecular Structure of a Left-handed Double Helical DNA Fragment at Atomic Resolution, Nature, 171, 737-738.

Watson G.N. (1966) A Treatise on the Theory of Bessel Functions, Cambridge University Press.

Watson J.D. (1976) The Molecular Biology of the Gene (3rd ed.), Benjamin, Calif.

Watson J.D. and Crick F.H.C. (1955) Molecular Structure of Nucleic Acids: A Structure for Deoxyribonucleic Acid, Nature, 171, 737-738.

Weston V.H. (1978) Electromagnetic Inverse Problem, in Uslenghi L.E. (ed.), Electromagnetic Scattering, Academic Press, N.Y., 289-313.

Wing R., Drew H., Takano T., Broka C., Tanaka S., Itakura K. and Dickerson R.E. (1980) Crystal Structure Analysis of a Complete Turn of B-DNA, Nature, 287, 755-758.

Wit A.L. and Cranefield P.F. (1978) Reentrant Excitation as a Cause of Cardiac Arrhythmias, Am. J. Physiol., 235, H1-H17.

Woolfson M.M. (1961) Direct Methods in Crystallography, Clarendon Press, Oxford.

Wu D., Denes P., Dhingra R.C., Wyndham C.R. and Rosen K.M. (1976) Quantification of Concealed Conduction Utilizing $S_1S_2S_3$ Stimulation, Circ. Res., 39, 659-665.

Zimmerman S.B. and Pheiffer B.H. (1979) Helical Parameters of DNA do not Change when DNA Fibres are Wetted: X-ray Diffraction Study, Proc. Nat. Acad. Sci. USA, 76, 2703-2707.